REGISTERED No. M. 3121

PROCEEDINGS

OF THE

INDIAN ACADEMY OF SCIENCES

VOL. XLI] *SECTION A* [No. 4

UNIVERSITY OF MICHIGAN
JUL 6 1955
PHYSICS LIBRARY

APRIL 1955

Price Rs. 2 or 3 Sh. *Annual Subscription Rs. 18*

THE STRUCTURE AND OPTICAL BEHAVIOUR OF JADEITE

By Sir C. V. Raman and A. Jayaraman
(*Memoir No. 70 of the Raman Research Institute, Bangalore*)

Received March 16, 1955

I. Introduction

Jadeite is a mineral to which much interest attaches by reason of the use which has been made of it since ancient times—especially in China—for the fabrication of works of art which have been greatly admired. It is a much rarer material than nephrite which is another mineral product to which the name jade is also commonly given; happily, the two substances are readily distinguishable from each other. In its finest forms, jadeite is known as gem-jade or imperial jade and differs from the commoner varieties in its optical characters; it exhibits a beautiful green colour resembling that of emerald and is semi-transparent instead of being cloudy or opaque. There is an extensive literature which deals with jadeite from the mineralogical point of view; works on gemmology also naturally devote a good deal of space to it. We have not however found in the literature any explanation of the optical characters which jadeite exhibits in terms of its structure. Indeed, it does not appear to have been realised that there is a problem awaiting answer in that connection. The opportunity for examining the subject presented itself to us when a collection of specimens of Burmese jade which was acquired for the Museum of this Institute became available for our studies.

2. The Structure of Jadeite

The chemical composition of jadeite corresponds to the formula $NaAl(SiO_3)_2$ and indicates that it belongs to the series of monoclinic pyroxenes of which diopside $CaMg(SiO_3)_2$ is a typical example. X-ray studies[1, 2, 3] support the latter conclusion, the X-ray powder diagrams of jadeite and diopside showing a close resemblance to each other. The published chemical analyses[4] show the presence in jadeite of varying proportions of other oxides, notably the oxides of calcium and magnesium, thereby indicating that diopside is present as a minor constituent in the mineral. Physically, jadeite is a polycrystalline aggregate. Examination of thin sections under the polarising microscope reveals the presence of individual crystals which

are birefringent, their greatest and the least refractive indices being 1·667 and 1·654 respectively.

When light enters a polycrystalline aggregate in which the individual grains are birefringent, it would necessarily suffer reflection and refraction at the intercrystalline boundaries. The smaller the individual crystallites, the more numerous would be such reflections and refractions. Hence, on this view, none of the incident light would succeed in penetrating through the material, if it is fine-grained and also of sufficient thickness. In the absence of any absorption properly so-called, the incident light would mostly be sent back towards the source by repeated internal reflections. In other words, a fine-grained aggregate of birefringent crystals should present a high albedo by reflected light and appear opaque in transmission.

The foregoing inferences based on geometrical optics do not however represent what is actually observed in the case of jadeite. Indeed, the converse is much nearer the truth; in other words, the specimens exhibiting a coarse structure are those which are opaque to light and reflect it strongly, while the specimens which are most readily penetrated by light are those having a fine-grained structure. Why this is so is the problem considered in the present paper.

3. THE X-RAY DIFFRACTION PATTERNS

To establish the relationship between structure and optical characters briefly indicated above, we have recorded the X-ray diffraction patterns of specimens exhibiting diverse optical behaviour. Three of these diagrams which are representative of the rest are reproduced in Plate XII. The film to specimen distance was in every case 5 cm. and MO$K\alpha$ radiation was employed to record the patterns.

Figure 1 was recorded with a piece of jade about half a millimetre thick cut out from one of our specimens and polished flat on both sides. The X-ray pencil passed normally through the specimen. The plate exhibited a notable measure of translucency; when placed on a printed sheet of paper, the finest print could easily be read through it. The X-ray record shows clearly that the crystallites present in the specimen are of varying sizes and that crystallites of very small sizes are the most numerous; though individual bright spots appear in the pattern, there are also numerous fainter ones which run into each other and appear as continuous rings.

Figure 2 is an X-ray diagram taken with similar arrangements, the X-ray pencil now grazing the edge of a flat ellipsoidal piece of green jade. It will be seen that the pattern is of the same general nature as that reproduced

in Fig. 1 except that the continuity of the rings is less evident and the spottiness more conspicuous.

Figure 3 is the pattern recorded with a greenish-white specimen which was nearly opaque to light. It will be seen that the pattern consists of a number of individual spots, some of which are very bright. But hardly anything in the nature of continuous rings can be discerned. In other words, the specimen consists of relatively large crystallites.

4. Theoretical Considerations

To understand the optical behaviour of jadeite, we have to put aside the considerations based on geometrical optics and deal with the matter from the standpoint of the wave theory of light. It is useful, in the first instance, to think of an extreme case in which the individuals in the polycrystalline aggregate are assumed to be of very small dimensions even in comparison with the wave-length of light and are also randomly orientated. Such an aggregate could obviously be regarded as equivalent to an optically isotropic medium with a refractive index which is the average of the principal refractive indices of the crystallites. A beam of light incident normally on a slab of such material would traverse it as a coherent wave-train, and emerge from the rear surface of the slab, but not the whole of the incident energy would succeed in so emerging. For, in consequence of the varying orientations of the crystallites and their birefringence, the medium would exhibit local variations in optical behaviour, and there would therefore be a diffusion of light both forwards and backwards. In the limiting case when the grains are very small, the diffusion would extend approximately with equal strength both forwards and backwards. More generally, however, there would be a concentration of the diffused radiation in the forward directions surrounding the path of the regularly transmitted light-beam. There would also be a diffusion of light backwards towards the light-source, but this would be less conspicuous.

The situation set forth above would be modified if the grains of the medium are of dimensions comparable with the wave-length of light. Nevertheless, provided the birefringence of the individual crystallites is small, we may continue to regard the aggregate as an isotropic medium which exhibits local variations in refractive index depending on the direction of vibration in the light beam traversing the material. These variations would result in the phase of light-waves passing through an individual grain being either accelerated or retarded as the case may be. The errors in phase thus resulting when summed up over the total path would tend to cancel each other out; the residual errors in phase would be the smaller relatively to

the total path, the more numerous the particles are, in other words the smaller is their size. Thus, it follows that if the birefringence is not large and the material is very fine-grained, the waves of light would succeed in penetrating the medium. Nevertheless, the residual errors in phase would result in a reduction of the absolute intensity of the transmitted beam and ultimately also its total extinction. The missing energy would appear for the most part in a cone of diffuse radiation surrounding the direction of propagation of the incident light-beam.

Wave-theoretical considerations also indicate that the proportion of the incident energy sent back towards the source as a diffuse reflection by a fine-grained aggregate would be far less than that indicated by considerations based on geometrical optics. For, the phase of a reflected wave relatively to the incident wave depends on whether the first or the second medium on the two sides of the reflecting boundary has the higher index. The boundaries encountered by a light-wave traversing a polycrystalline aggregate would obviously be of both kinds and in approximately equal numbers. The reflection in the two cases being in opposite phases, their effects would tend to cancel each other more or less completely in all directions.

Thus we may sum up the situation by the statement that light would penetrate into a polycrystalline aggregate much more freely and would be reflected backwards much less freely than that might be anticipated in terms of geometrical optics, provided that the aggregate is sufficiently fine-grained. Thus, we have an explanation in general terms of the semi-transparency exhibited by the finest varieties of jadeite.

5. SUMMARY

The phenomenon of semi-transparency exhibited by the finer varieties of jadeite is discussed. X-ray diffraction studies and theoretical considerations based on wave-optics alike indicate that this is a consequence of the very small particle size of the crystallites forming the aggregate.

6. REFERENCES

1. Wyckoff, R. W. G., Mervin, H. E. and Washington, H. S. — *Am. Jour. Sci.*, 5th Series, 1925, **10**, 383.
2. Warren, B. E. and Biscoe, J. .. *Zeit. F. Krist.*, 1931, **80**, 391.
3. Yoder, H. S. .. *Am. Jour. Sci.*, 1950, **248**, 225.
4. Hintze Carl .. *Handbuch Der Mineralogie, Zweiter Band*, 1897, p. 1175.

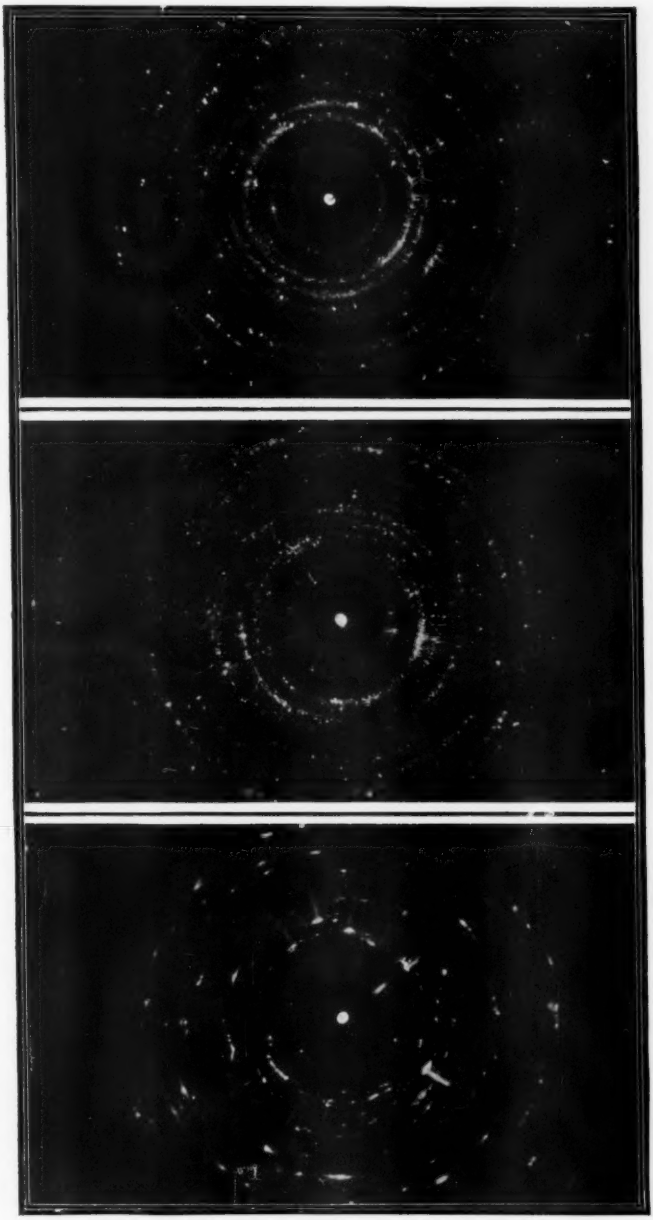

FIG. 1

FIG. 2

FIG. 3

X-ray Diffraction Patterns of Jadeite

RAMAN SPECTRUM OF CRYSTALLINE CADMIUM SULPHATE

By T. S. Krishnan and P. S. Narayanan

(From the Department of Physics, Indian Institute of Science, Bangalore-3)

Received March 4, 1955

(Communicated by Prof. R. S. Krishnan, F.A.SC.)

1. Introduction

Cadmium sulphate ($3CdSO_4.8H_2O$) differs from most other crystalline sulphates containing water of crystallisation in that the proportion of water is very much less and the crystal structure analysis (Lipson, 1936) shows that the forces between the cadmium and SO_4 ions are so large that to a first approximation one can regard the structure as composed of Cd and SO_4 groups in which the water molecules occupy the interstices. Although the spectrum of its water of crystallisation has been studied and compared with those of other crystals containing water of crystallisation (Hibben, 1937; P. G. N. Nayar, 1938) its Raman spectrum has not been so far investigated in any great detail. Its Raman spectrum recorded by Hibben and reproduced in his book ' Raman Effect and its Chemical Applications " (1939) does not show much detail and Hibben limited himself to pointing out the splitting in the crystalline state of the triply degenerate oscillation (1106 $cm.^{-1}$) of the free SO_4 ion and no further study of the spectrum was made by him. Besides these, the strongest Raman line 1003 $cm.^{-1}$ alone was recorded by Embirikos (1930) and compared with those of other sulphates. An investigation of its Raman spectrum made by us recently revealed such an abundance of detail due to the use of spectrographs of large dispersion, improved Rasetti technique and more efficient filtering of the exciting radiation that it was felt desirable to report the results given below.

2. Experimental Details

The crystals used were prepared from the aqueous solution of the salt by the method of slow evaporation and were about 1 cm. $\times$ 5 mm. $\times$ 5 mm. in size with well-developed faces. Besides those grown in our laboratory, one other fine specimen also grown from solution by the same method was made available to us through the kindness of Prof. J. P. Mathieu to whom our sincere thanks are due.

The resonance radiation of mercury (λ 2537) from a specially designed water-cooled quartz arc placed in the field of an electromagnet was used to excite the Raman spectrum. Using the medium quartz spectrograph and slit width of ·025 mm. exposures of the order of 2 hours were found to be sufficient to record the complete spectrum. It is worth mentioning here that though the experimental arrangement was found to be adequate to record the symmetric oscillation (1006 cm.$^{-1}$) of the SO_4 ion and the water-bands in a couple of minutes, even exposures of the order of 2 days failed to bring out the second order lines. In order to study the lattice spectrum in more detail a large quartz spectrograph with a dispersion of about 40 cm.$^{-1}$ per mm. in the λ 2537 region was employed. With this instrument using a slit width of ·05 mm. an exposure of the order of a day was enough to record the complete Raman spectrum.

3. Results and Discussion

The Raman spectrum of cadmium sulphate $3CdSO_4 \cdot 8H_2O$ consists of 23 frequency shifts all of which except the waterbands are marked in Fig. 1 *b* of Plate XIII. Fig. 1 (*a*) is the spectrum taken with the large quartz spectrograph and enlarged to the same extent as its microphotometer record shown in Fig. 1 (*b*). The waterbands as well as the other Raman lines could also be easily identified in Fig. 2 *a* which is an enlargement of the spectrogram taken with the smaller instrument while Fig. 2 *b* shows the direct mercury spectrum. The observed frequency shifts are 45 (2), 56 (7), 64 (7), 75 (2), 89 (2), 104 (5), 124 (4), 138 (5), 170 (2), 190 (3), 219 (5), 261 (1), 331 (3), 456 (6), 462 (5), 626 (2), 1006 (30), 1077 (3), 1101 (9), 1115 (10), 1157 (3), 3347 (7) and 3428 (15) cm.$^{-1}$ the figures within brackets representing the visually estimated relative intensities. The values of the frequency shifts of all the lines except those of the waterbands were determined from measurements made on the spectrogram taken with the El (large) spectrograph and may therefore be regarded as accurate. We will now consider the observed spectrum in relation to the known crystal structure.

The crystal structure of $3CdSO_4 \cdot 8H_2O$ has been determined by Lipson (1936) and reveals certain interesting features. Unlike most other sulphates with water of crystallisation, like alum, $KAl(SO_4)_2\,12 \cdot H_2O$ or Epsom salt $MgSO_4 \cdot 7H_2O$ where the water is known to form an essential part of the structure, in cadmium sulphate eight-third hydrate the main structure is governed by the forces between the metal atoms and the SO_4 groups and the water molecules just fill up the interstices. The crystal belongs to the monoclinic space group $I\,2/h$ (C^6_{2h}) and contains four molecules of $3CdSO_4 \cdot 8H_2O$ in the unit cell. In the crystal, only 4 Cd and 4 SO_4 occupy

special positions and are located on the two-fold axis and all the remaining atoms are in general positions. The local symmetry of even the special SO_4 group in the crystal is only C_i and as such it is only reasonable to suppose that on account of this and also the known multiplicity of ions in the unit cell, the degeneracies of the frequencies belonging to tetrahedral SO_4 ion are removed. Actually one observes only 8 frequencies which could be attributed to the internal oscillations of the SO_4 ion as shown below.

RAMAN FREQUENCIES OF THE SO_4 ION IN cm.$^{-1}$

In the free state →	454	622	983	1106
In $3CdSO_4.8H_2O$ →	456, 462	626	1006	1076, 1101, 1115, 1157

The Raman line 1076 cm.$^{-1}$ has been explained by Hibben as a combination of the frequencies 450 and 626 cm.$^{-1}$; this appears to be erroneous since in the crystal the line 1076 cm.$^{-1}$ is very intense while the fundamental 626 cm.$^{-1}$ itself is very weak. While the degeneracy of the oscillation (1106 cm.$^{-1}$) of the free SO_4 ion is seen to be removed completely and splitting is evident, one notices, however, a different behaviour of the other two degenerate oscillations. In the crystal the Raman lines due to the splitting of the doubly degenerate oscillation are neither well resolved nor sharp as for example in gypsum which also belongs to the monoclinic class and is known to contain loosely-bound water of crystallisation (Wooster, 1936). Also in the present study only a very weak line corresponding to the triply degenerate oscillation of the free ion is observed at 626 cm.$^{-1}$ while in gypsum three fairly sharp lines (618, 622 and 672 cm.$^{-1}$) have been recorded by R. S. Krishnan (1945) in the same region. Further the spectrum of the water of crystallisation is comparatively simple and consists of only two bands the frequency shifts of which agree well with those reported by Hibben (1937) and by P. G. N. Nayar (1938). These observed features of the lines due to the internal oscillations of the SO_4 ions and of the water of crystallisation suggest that in the crystalline state while the influence of the additional forces due to the neighbouring ions is by no means negligible, it does not lead to any major distortion of the disposition of the bonds from their idealized nature as expected by Lipson (*loc. cit.*).

Besides the 8 Raman lines due to the internal oscillations of the SO_4 groups and the two waterbands, the Raman spectrum contains 13 low frequency shift lines which arise from lattice oscillations. Neglecting the water molecules, a simple group theoretical calculation shows that the number of translatory type of oscillations permitted in Raman effect is 36, of which

16 are of type A_{1g} and the remaining 20 of type B_{1g}. However, we find that even under the high dispersion of the large quartz spectrograph the crystal exhibits only 13 Raman lines which are attributable to the translatory type of lattice oscillations. Since it is known that for this crystal only the vibrations of type A_{1g} will show a variation of intensity with orientation an attempt was made to study the variation of the intensity with orientation with incident unpolarised light. The observations indicate that the lines 219, 104, 138 cm.$^{-1}$ may tentatively be assigned to type A_{1g} and those with frequency shifts 170, 190, 331 cm.$^{-1}$ to the type B_{1g}. It was also observed that among the Raman lines due to the SO_4 ions and the waterbands those at 462 and 3347 are of type A_{1g}.

It is a pleasure for us to thank Professor R. S. Krishnan, for his help and encouragement.

SUMMARY

The Raman spectrum of a single crystal of cadmium sulphate ($3CdSO_4.8H_2O$) has been investigated using the resonance radiation of mercury as exciter. The recorded spectrum reveals 23 Raman lines and of these 8 have been attributed to the internal oscillations of the SO_4 ion and 2 to the oscillations of the water molecules. While the number of translatory type of lattice oscillations permitted by the selection rules is 36, only 13 are observed experimentally. A tentative assignment for some of these vibrations has been made from a study of the variation of intensity of the lines with orientation.

REFERENCES

1. Embirikos, N. .. *Zeit. Phys.*, 1930, **65**, 266.
2. Hibben, J. H. .. *Journ. Chem. Phys.*, 1937, **5**, 166.
3. ———— .. *The Raman Effect and its Chemical Applications*, 1939, 332, 399.
4. Krishnan, R. S. .. *Proc. Ind. Acad. Sci.*, 1945, **22A**, 274.
5. Lipson, H. .. *Proc. Roy. Soc.*, London, 1936, **156A**, 462.
6. Nayar, P. G. N. .. *Proc. Ind. Acad. Sci.*, 1938, **8A**, 419.
7. Wooster, W. A. .. *Zeits. f. Krist.*, 1936, **94**, 376.

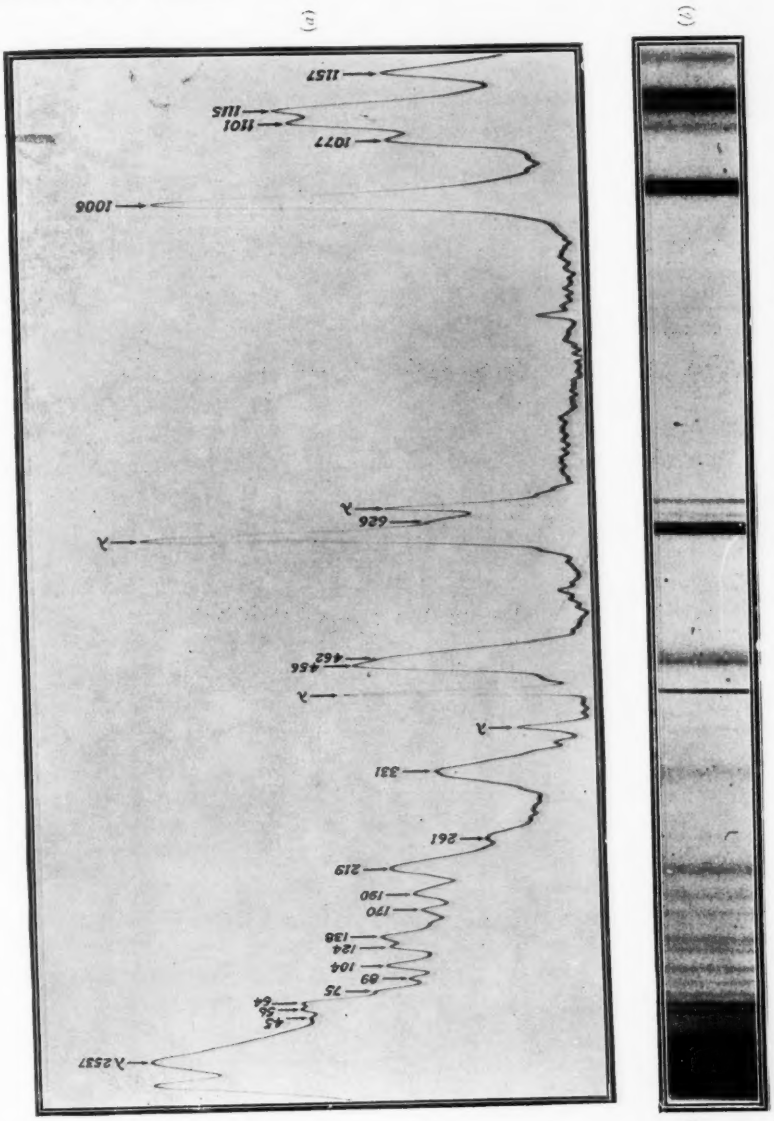

FIG. 1. (*a*) Raman spectrum of crystalline cadmium sulphate taken with a large quartz spectrograph
(*b*) Its microphotometer record

FIG. 2. (*a*) Raman spectrum of crystalline cadmium sulphate taken with the medium quartz spectrograph.
(*b*) Direct Mercury Spectrum.

RELAXATION EFFECTS IN AQUEOUS SOLUTIONS OF SULPHUR DIOXIDE

By M. Krishnamurthi

(From the Physical Laboratories, Osmania University, Hyderabad)

Received March 4, 1955
(Communicated by Dr. S. Bhagavantam)

1. Introduction

Sound absorption in most liquids is proportional to the square of frequency as required by the classical formulæ for viscosity and heat conduction losses. Departures from this law are now generally ascribed to relaxation phenomena in intermolecular processes. Herzfeld and Rice (1928) followed by Kneser (1948), Markham (1950), Liebermann (1949), Hall (1949) and Lamb and Pinkerton (1949) have offered explanations based on relaxation. Herzfeld and Rice (*loc. cit.*) attributed anomalous absorption to what may be described as thermal relaxation, that is a lag in the adjustment of energy between the external and internal degrees of freedom of the individual molecules during passage of a sound wave. Kneser, Markham and others have treated the problem more generally and visualised other types of relaxation. In associated liquids there may be changes in the degree of association in the presence of a sound wave, which may take a finite time and hence lead to absorption of energy. Liebermann postulated a disturbance in the dynamic equilibrium between the solute and its ions in an aqueous solution of $MgSO_4$, to explain the observed peak in absorption at a particular frequency in such solutions. Whatever may be the relaxation mechanism proposed, some common conclusions follow from all these theories. One is that there is a frequency ω_0, called the relaxation frequency, and an associated relaxation time τ characteristic of each liquid. In regions not too close to this frequency, absorption is linearly related to the square, and velocity is independent, of frequency. On the other hand, near this frequency, velocity shows dispersion. At a frequency which is a simple function of ω_0 and less than the latter, the absorption per wave-length ($\alpha\lambda$) goes through a peak when plotted against frequency. Sulphur dioxide, ammonia and carbon dioxide have all been studied thoroughly in their gaseous states and are known to exhibit the abovementioned relaxation phenomena. It is considered interesting to see how they behave in solution. In the present work, solutions of these gases in some typical liquids are studied for the first time.

2. EXPERIMENTAL METHOD

Employing a method similar to that described by Pellam and Galt (1946), in which instead of a reflector, a second crystal is used as the receiver, sound absorption measurements have been made in several gaseous solutions. The two crystals used have identical fundamental frequencies of 2·12 Mc/s and are worked at their 3rd, 5th and 7th harmonics besides their fundamental. The temperature is controlled to within $\pm$ 0·5° C. The alignment of the crystals as well as the accuracy of the method as a whole is tested by taking measurements in double distilled and suspension—free water at all the frequencies. For attenuation measurements, the height of the received pulse is adjusted to be constant by using a calibrated standard decade attenuator with a bandwidth of 0 to 30 Mc/s and capable of reading correct to 0·1 db. The attenuation measurements are accurate to within 3 to 6 per cent., the lower limit applying to the higher frequencies.

Sulphur dioxide used in this work is manufactured by I.C.I. as a refrigerant and is of guaranteed purity. Ammonia is obtained as an aqueous solution (sp. gr. 0·88) of analytical purity. Carbon dioxide is a commercial sample obtained locally. Sulphur dioxide solutions are estimated by titration in the usual way. Strengths of ammonia solutions are determined from density data. Only a saturated solution of carbon dioxide has been studied and no attempt is made to determine its strength.

3. RESULTS AND DISCUSSION

The highest concentrations of sulphur dioxide studied are 0·1 M in water, 0·55 M in Benzene, 0·95 M in Ethyl Alcohol and 0·3 M in Nitrobenzene. Ammonia is dissolved only in water and the maximum concentration studied is 0·55 M. All these solutions are investigated at the room temperature. Carbon dioxide which is very sparingly soluble is examined at the low temperature of 8° C. so that more of the gas will dissolve in water. A representative set of results for all these solutions are presented in Figs. 1 and 2.

From the above graphs it is seen that solutions of carbon dioxide and ammonia in water behave normally in that $\alpha\lambda$ increases linearly with frequency at all concentrations. Therefore solutions of these gases in other solvents are not studied. On the other hand, aqueous sulphur dioxide solutions show a peculiar behaviour, in several ways. The attenuation increases very rapidly with concentration. The plot of $\alpha\lambda$ against frequency is not a straight line with a positive slope. It actually exhibits peaks which appear to occur at lower frequencies for lower concentrations. At the higher

concentrations, the frequency of peak absorption ($\alpha\lambda$) reaches a limit and becomes independent of concentration. All these peculiarities point to a relaxation phenomenon and the characteristic relaxation frequency at the higher concentrations lies in the range 2 to 15 Mc/s while at lower concentrations it is less than 2 Mc/s. As a probable cause for this phenomenon, one can think of the gas itself or the chemical reaction going on in the aqueous solutions. That the gas is not the cause of this phenomenon follows from the fact that non-aqueous solutions of sulphur dioxide do not show any such relaxation. This leaves the dynamic chemical equilibrium between H_2SO_3 and H^+ and HSO_3^- ions, namely $H_2SO_3 \leftrightharpoons H^+ + HSO_3^-$,

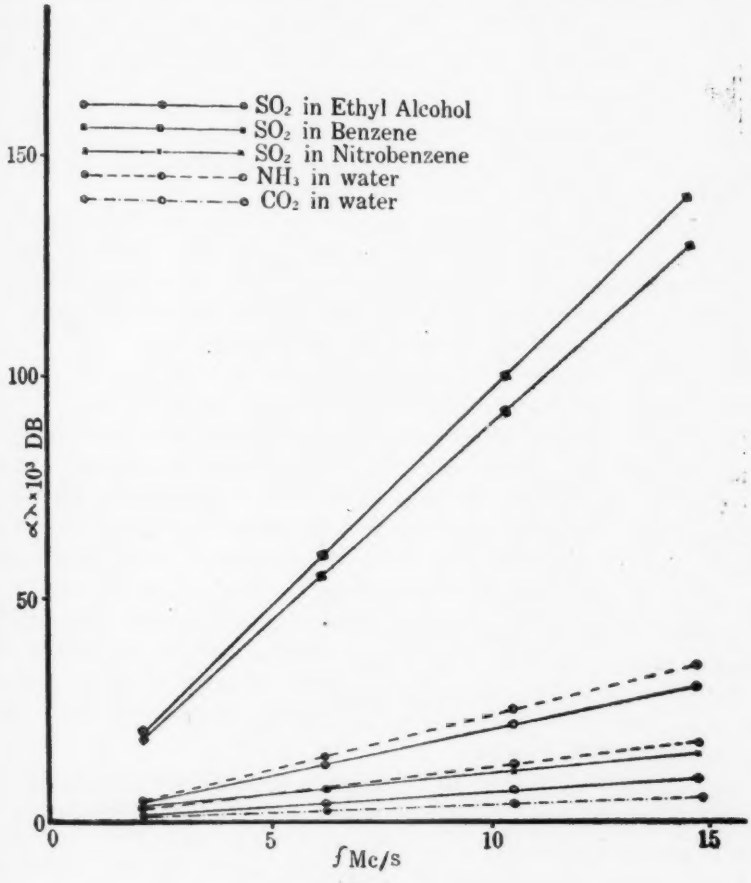

Fig. 1

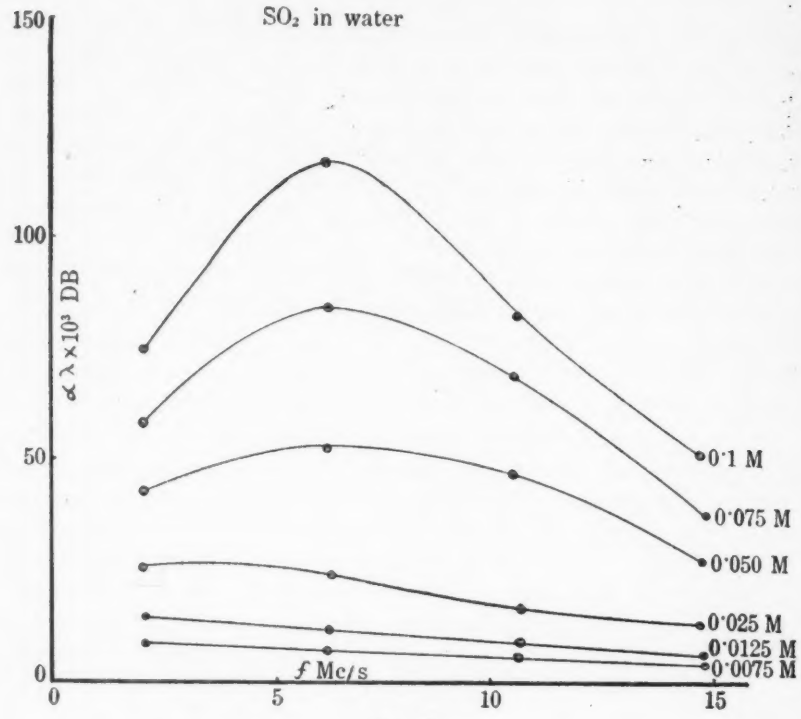

FIG. 2

as the probable cause. In the presence of the ultrasonic beam, especially n the particular range of frequencies used in this work, there seems to be a change in this equilibrium followed by absorption of energy. This means that the above reaction is proceeding in either direction at a rate comparable with the period of oscillation of the sound waves and that with increasing concentration it is increasing upto a limit.

Studies in aqueous solutions of bisulphites of potassium and sodium as well as sodium sulphite show similar results. This confirms that where an HSO_3 ion is concerned, relaxation phenomena are again exhibited. Details of these investigations, which are of a confirmatory type, will be published separately.

It is, however, necessary to explain why aqueous solutions of ammonia and carbon dioxide which also undergo chemical changes in solution, do not exhibit relaxation effects. These three gases have been thoroughly studied by Fricke (1939) and Petralia (1953) among others and they have found

that carbon dioxide exhibits a peak in absorption at about 20 Kc/s. The corresponding values for sulphur dioxide and ammonia are 1·040 Mc/s. and 393 Mc/s. respectively. The relaxation times in the latter have been calculated as $1{\cdot}85\times10^{-7}$ sec. and $8{\cdot}1\times10^{-10}$ sec. respectively. It means that these are the times in which the extra translational energy supplied by the sound beam is converted by the respective molecules into internal energy. It is reasonable to expect that the conversion of chemical energy into kinetic energy or back also involves times of this order of magnitude for the two types of molecules. Thus, the frequency range to which the present studies are confined explains the occurrence of a relaxation phenomenon in the case of sulphur dioxide solutions only and not in ammonia solutions. On the same basis, the rate of reaction in the case of carbon dioxide will be so slow that so far as the frequencies involved in the present experiments are concerned, it is as good as not taking place at all.

4. Summary

Absorption of ultrasonic energy by solutions of sulphur dioxide ammonia and carbon dioxide in water, and in the case of the first gas in several other organic solvents, has been studied by the pulse technique. Aqueous sulphur dioxide solutions exhibit a relaxation phenomenon in the frequency range of 2 to 15 Mc/s. as evidenced by the absorption ($\alpha\lambda$) showing a peak in this region when plotted against frequency. This peak occurs at higher frequencies for higher concentrations upto a limit after which the peak absorption frequency does not change with concentration. These observations, along with the fact that no such effect is observed in the case of other solvents, lead to the conclusion that the phenomenon is caused by a change in the dynamic equilibrium between H_2SO_3 molecules and H^+ and HSO_3^- ions in the presence of the sound beam. The absence of a similar effect in the case of ammonia and carbon dioxide solutions which also involve a chemical reaction is explained on the basis of the widely differing relaxation times of the three gases.

In conclusion, the author acknowledges with grateful thanks the guidance and help given by Professor S. Bhagavantam who suggested this problem.

References

1. Hail .. *Phys. Rev.*, 1948, **73**, 775.
2. Herzfeld and Rice .. *Ibid.*, 1928, **31**, 691.
3. Kneser .. *Ergeb. Exakt. Naturwiss.*, 1949, **22**, 121.
4. Lamb and Pinkerton .. *Proc. Roy. Soc.*, 1949, **76**, 1520.
5. Liebermann .. *Phys. Rev.*, 1949, **76**, 1520.
6. Markham .. *J. Acous. Soc. Am.*, 1950, **22**. 628.
7. Pellam and Galt .. *J. Chem. Phys.*, 1946, **14**, 608.

ACHROMATIC COMBINATIONS OF BIREFRINGENT PLATES

Part I. An Achromatic Circular Polarizer

By S. Pancharatnam

(*Memoir No.* 71 *of the Raman Research Institute, Bangalore*)

Received March 5, 1955

(Communicated by Sir C. V. Raman)

1. Introduction

When polarised light of any form is incident on a birefringent plate, the phase retardation δ introduced between the two waves during their passage through the plate, is far from being the same for all wave-lengths; in fact, being given by $\delta = (2\pi/\lambda)(\mu_1 - \mu_2)\, t$, this retardation of phase varies almost inversely as the wave-length λ, that is, if the birefringence $(\mu_1 - \mu_2)$ does not disperse notably with wave-length. So much so, that what acts as a quarter-wave retardation plate for the deep red end of the spectrum will, for the wave-lengths in the deep violet, behave practically as a half-wave plate.

The question of the achromatisation of devices in this field has engaged the attention of several workers. References to their investigations may be found cited in the paper entitled, "*Réalisation d'un quart d'onde quasi-achromatique par juxtaposition de deux lames crystallines de même nature*" by Destriau and Prouteau[1]; the two plates referred to, do not, of course, have their principal planes parallel, and the 'compound plate', comprising the two superposed plates, can transform incident circularly polarised light to plane polarised light vibrating at a certain azimuth, or *vice-versa*. But the device described by these authors would more properly be called an achromatic circular polarizer (or analyser); to call it an achromatic quarter-wave plate would be incorrect since the combination does not have the usual at ributes of a quarter-wave plate—it cannot, for example, be used for the analysis of elliptically polarised light in the usual manner of an ordinary $\lambda/4$ retardation plate.

Nevertheless it is as a circular polarizer or analyser that a quarter-wave plate is often used—as when it is inserted with its principal planes at an angle of 45° to those of a nicol, in a petrographic microscope. And we shall reserve for the second part, the problem of superposing birefringent plates in such a manner that the combination as a whole behaves as an achromatic

quarter-wave plate. In this paper we shall describe a circular polarizer which is expected to have a much higher degree of achromatism than the one discussed in the paper quoted above. The description of the device may be found at the end of the next section; its applicability is not limited to the visible. The birefringent plates in our discussion are all of the same material; the dispersion of the birefringence of this material is not assumed to be negligible—though such an assumption would indeed be justified for plates of muscovite mica, at least for the visible and ultra-violet wave-lengths.[2]

2. The Achromatic Circular Polarizer

The Poincaré sphere* lends itself very conveniently for the theoretical discussion of the two problems at hand—especially the one to be treated later. (As is well known, any point on the surface of this sphere, of latitude 2ω and longitude $2l$, represents an elliptic vibration of ellipticity $|\tan \omega|$ the major axis of which makes an angle l with a fixed reference direction. Passage through a birefringent plate of retardation δ, the orientation of whose principal planes are given by O and O′, is equivalent to the operation of rotating the sphere by an angle δ about the equatorial diameter OO′.)

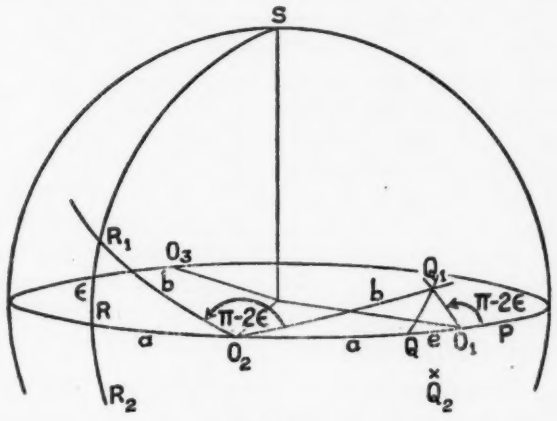

Fig. 1

A quarter-wave plate with its fast vibration direction in the orientation O_3 (Fig. 1) can transform the incident linear vibration R—inclined at 45° to the principal planes—to circularly polarised light represented by the pole S—but only for the wave-length λ for which its retardation is exactly

* Expositions in English, of the properties of the Poincaré sphere are apparently not so common in text-books, but may be found in numerous recent articles: see *e.g.*, Ramachandran and Ramaseshan, *J. Opt. Soc. Am.*, 1952, **42**, 49.

$\pi/2$. For the wave-lengths λ_1 and λ_2 on either side of λ, for which the respective retardations are $\pi/2 - \epsilon$ and $\pi/2 + \epsilon$, the emergent light will be circularly polarised only if the incident vibrations are given respectively by R_1 and R_2; these two points lie on the meridian SR and are equidistant from R (arc $R_1R = \epsilon$). Thus if the ellipticity of the incident vibrations were dispersed roughly along R_1R_2 in a proper fashion, the quarter-wave plate would transform the vibrations for all the wave-lengths to approximately circular polarised light.

The required dispersion of the ellipticity of the vibrations incident on the quarter-wave plate, it may be seen, can be effected by allowing plane polarised light P to first pass successively through two half-wave plates, the orientation of whose fast axes are given by O_1 and O_2 respectively and are to be determined. In order that the light emerging from the third retardation plate (*i.e.*, the quarter-wave plate) should be exactly circularly polarised for all the three wave-lengths λ_1, λ and λ_2, the state of polarization of the light incident on the second half-wave plate for these three wave-lengths should be given by Q_1, Q and Q_2 respectively—got by the following construction. The point Q is marked off on the equator such that $O_2Q = O_2R$ ($= a$ say); the arc O_2Q_1 is drawn such that $R_1O_2Q_1 = \pi - 2\epsilon$, and $O_2Q_1 = O_2R_1$ ($= b$ say); Q_2 is a point symmetrically placed with respect to Q_1 on the lower hemisphere. The first half-wave plate has to be oriented in a position depending on the quantity a which determines the orientation of the second; for we must have $O_1Q_1 = O_1Q_2 = O_1Q$ ($= e$ say); and further we must choose a such that $P\hat{O}_1Q_1 = \pi - 2\epsilon$. If the plane of the incident vibration P be now set such that $O_1P = e$, then the light emerging from the combination will be circularly polarized for the three wave-lengths λ_1, λ and λ_2 for which the retardations of a quarter-wave plate would be $\pi/2 - \epsilon$, $\pi/2$, and $\pi/2 + \epsilon$ respectively; ϵ is an arbitrarily chosen parameter on which will depend the two opposing characteristics of the combination: the range and the degree of its achromatism.

By the use of spherical trigonometry we shall get an equation for a involving the parameter ϵ. Thus, from the equilateral triangle Q_1O_1Q, denoting arc Q_1Q by c

$$\cos 2\epsilon = -\cos^2 Q_1QO_1 + \sin^2 Q_1QO_1 \cos c$$

or

$$1 + \cos 2\epsilon = 2\sin^2 Q_1QO_1 \cos^2 c/2$$

From the triangle Q_1O_2Q we have, writing $Q_1O_2Q = C$,

$$\frac{\sin c}{\sin C} = \frac{\sin b}{\sin Q_1QO_1}$$

i.e.,

$$\sin Q_1QO_1 \cos \frac{c}{2} = \frac{\sin b \sin C}{2 \sin \frac{c}{2}}$$

We thus obtain

$$\sin^2 b \sin^2 C = 2 \sin^2 \frac{c}{2} (1 + \cos 2\epsilon) \tag{1}$$

We shall express the quantities on the left-hand side in terms of a. Denoting R_1O_2R by C'

$$\sin C = \sin (2\epsilon - C') = \sin 2\epsilon \cos C' - \cos 2\epsilon \sin C' \tag{2}$$

From the right-angled triangle R_1O_2R

$$\sin C' = \frac{\sin \epsilon}{\sin b}$$

and

$$\cos C' = \frac{\cos \epsilon - \cos a \cos b}{\sin a \sin b} = \cos \epsilon \cdot \frac{\sin a}{\sin b}$$

since

$$\cos b = \cos \epsilon \cos a \tag{3}$$

Substituting the expressions obtained for $\cos C'$ and $\sin C'$ in (2)

$$\sin C = \frac{\sin \epsilon}{\sin b} (2 \cos^2 \epsilon \sin a - \cos 2\epsilon)$$

Introducing this in (1)

$$(2 \cos^2 \epsilon \sin a - \cos 2\epsilon)^2 = 2 \cot^2 \epsilon (1 - \cos c) \tag{4}$$

It remains to express $\cos c$ in terms of a. From triangle Q_1QO_2

$$\cos c = \cos a \cos b + \sin a \sin b \cos C$$

$$\cos C = \cos (2\epsilon - C') = \cos 2\epsilon \cos C' + \sin 2\epsilon \sin C'$$

$$= \frac{\cos \epsilon}{\sin b} (\cos 2\epsilon \sin a + 2 \sin^2 \epsilon)$$

Usin this relation, together with (3), the expression for $\cos c$ becomes,

$$\cos c = \cos \epsilon (1 + 2 \sin^2 \epsilon \sin a - 2 \sin^2 \epsilon \sin^2 a)$$

Substituting for $\cos c$ in (4), and rearranging, we finally get an equation in $\sin a$ of the form:

$$A \sin^2 a + B \sin a + C = 0 \tag{I}$$

where

$$A = 4\cos^3\epsilon(1-\cos\epsilon)$$

$$B = 4\cos^2\epsilon(\cos 2\epsilon - \cos\epsilon)$$

$$C = 2\cot^2\epsilon(1-\cos\epsilon) - \cos^2 2\epsilon$$

An expression for e may be obtained from triangle Q_1O_1Q

$$\cos c = \cos^2 e + \sin^2 e \cos 2\epsilon$$

or

$$\sin^2 e = \frac{1-\cos c}{1-\cos 2\epsilon}$$

Using (4),

$$\sin e = \frac{2\cos^2\epsilon \sin a - \cos 2\epsilon}{2\cos\epsilon} \qquad \text{(II)}$$

The combination can produce exactly circularly polarized light for three wave-lengths. If we neglect the dispersion of the birefringence, the two extreme wave-lengths will obviously be $(1+f)\lambda$ and $(1-f)\lambda$—where we have denoted the ratio $\epsilon : \pi/2$ by f; the achromatism of the combination may however be considered to extend up to two wave-lengths *outside* this interval, for which the deviation from circular polarisation is roughly the same as the maximum deviation attained within this interval—which latter may be expected to occur in the neighbourhood of the wave-lengths $(1 \pm \cdot 5f)\lambda$. Thus we may with some arbitrariness, consider the combination to be achromatic within the range $(1 \pm 1\cdot 25f)\lambda$.

Choosing $2\epsilon = 47°$ (which means $1\cdot 25f = \cdot 325$), equation (I) when solved gives $\sin a = \cdot 6642$ or $a = 41°\,37'$.

The corresponding value of e obtained from (II) is $e = 13°\,43'$.

On the other hand, if we choose $2\epsilon = 36°$ (which means $1\cdot 25f = \cdot 25$) we obtain

$$a = 43°\,8';\quad e = 13°.$$

Thus to summarize our results, an achromatic arrangement for producing circularly polarised light is obtained by allowing parallel plane polarised light to pass normally through a combination of three superposed plates of the same material; the first two should be half-wave plates, and the last, a quarter-wave plate for the wave-length λ in the centre of the spectral range to be covered. Let θ be the angle by which the fast vibration direction of the first plate is turned with respect to the azimuth of the incident vibration; θ_1, the angle by which the fast axis of the second plate is turned with respect to the first; and θ_2 the corresponding angle between the fast axes of the

third and second plates—all the angles being measured in the same sense. The set of values

$$\theta_2 = \tfrac{1}{2}(\pi/2 + a) = 65° 49'; \quad \theta_1 = \tfrac{1}{2}(a + e) = 27° 40'; \quad \theta = \tfrac{1}{2} e = 6°52'$$

can be used for covering the range from $1 \cdot 325\ \lambda$ to $\cdot 675\ \lambda$ and is therefore suited for covering the entire visible spectrum if we choose $\lambda = 6000$ Å. A greater degree of achromatism can be attained at the expense of restricting the range to the major portion of the visible. Thus the set of values

$$\theta_2 = 66° 34'; \quad \theta_1 = 28° 4'; \quad \theta = 6° 30'$$

can be used for covering the range from $1 \cdot 25\ \lambda$ to $\cdot 75\ \lambda$. If the dispersion of the birefringence of the material used is not negligible, the values given above for the range over which the combination may be considered achromatic would have to be altered.

On turning the polarizing nicol to a perpendicular position, circularly polarized light of the opposite sense is produced. In fact in passage through any succession of 'elliptically birefringent' plates, two orthogonal states of the incident polarization will correspond to two orthogonal states of the emergent polarization. This is a consequence of a more general fact; the angular separation on the Poincaré sphere of two possible states of the incident polarization will be equal to the angular separation of the two corresponding states of the emergent polarization, since angular relationships remain invariant under any number of rotations. Thus if the nicol had been turned by 45°, the emergent wave-lengths would all have been practically plane polarized, but with the azimuths of the vibrations dispersed—an interesting illustration of the fact that the combination does not have the properties of a quarter-wave plate.

3. Experimental Verification

Two half-wave plates and a quarter-wave plate, all prepared of mica, were cemented together (using copal varnish) with their principal planes at inclinations slightly different from the first set of values given above—in order to cover an even wider range. The combination was laid on a mirror. With the aid of another mirror, white light from a point source was made to pass normally through a polaroid and the achromatic combination, and then back again—after which it reached the eye. As the polaroid is rotated a position is reached where the image of the source is completely extinguished; this is the position where a single passage through the compound plate would give circularly polarized light, while the double passage gives plane polarized light which is crossed out by the polarizer itself. When the experiment is

repeated with a single quarter-wave plate the image is never completely extinguished, but is highly coloured.

The two $\lambda/2$ retardation plates may first be cemented together at the proper angle θ_1, using the fact that the two together will rotate the plane of polarization of *any* incident linear vibration (of wave-length λ) by an angle $\pi - 2\theta_1$; (this can be proved by a construction given in Part II); the quarter-wave plate may then be cemented and adjusted, while the cement is still wet, to the position of best achromatism—as determined by the test given above.

SUMMARY

Circularly polarised light is obtained by superposing two half-wave plates and one quarter-wave plate, all of the same material, such that the fast vibration-directions of the successive plates make specific angles θ_1, θ_2, θ_3 with the azimuth of the linear vibration incident on the first plate. The required range of achromatism determines the optimum values of the angles. Thus, using mica retardation plates prepared for Na 5890, the range from 4000 Å to 7800 Å is covered with $\theta_1 = 6° 52'$, $\theta_2 = 34° 32'$, $\theta_3 = 100° 20'$; while the range 4400 Å to 7400 Å can be covered with superior achromatism by taking $\theta_1 = 6° 30'$, $\theta_2 = 34° 34'$, $\theta_3 = 101° 8'$.

REFERENCES

1. Destriau and Prouteau .. *J. Phys. et radium*, 1949, **10**, 53.
2. Mathieu .. *Bull. Soc. Franc. Min.*, 1934, **57**, 233.

ACHROMATIC COMBINATIONS OF BIREFRINGENT PLATES

Part II. An Achromatic Quarter-Wave Plate

By S. Pancharatnam

(*Memoir No. 71 of the Raman Research Institute, Bangalore*)

Received March 5, 1955

(Communicated by Sir C. V. Raman)

1. Introduction

A combination of birefringent plates which can be used for producing circularly polarised light over a fairly large range of wave-lengths has been described in Part I. And it has there been pointed out that the arrangement cannot be properly described as an achromatic quarter-wave plate. A quarter-wave plate when rotated between crossed nicols will show two extinction positions; the achromatic combination described will show none. And to understand under what conditions a combination of plates will have the properties of a single birefringent plate—let alone a plate which is in addition achromatic—it is desirable to have a general method for combining the action of a pile of birefringent plates when their principal planes are not parallel. We shall follow a method differing slightly from that which may be found described by Pockels.[1]

2. General Considerations

The problem of designing an achromatic elliptic polariser is the problem of transforming a particular state of polarization represented by P_1 (on the Poincaré sphere) to another particular state P_1'. But, of an achromatic retardation plate we require that it transforms *every* P to a corresponding P′ obtained by rotating the sphere by an angle δ about an equatorial diameter.

Passage of monochromatic polarised light through a succession of birefringent plates corresponds to a succession of rotations of the Poincaré sphere about a series of equatorial diameters. Any succession of rotations of the sphere can, however, be compounded into one single resultant rotation. Hence the action of a combination of birefringent plates corresponds to a rotation ϕ of the Poincaré sphere about some diameter EE′—which will in general be inclined to the equatorial plane; this means, of course, that the combination is equivalent to a single plate having elliptic birefringence—

E and E′ being the orthogonal elliptic vibrations propagated without change of form, and between which a phase retardation ϕ is incurred.

The condition that the combination should act as a purely birefringent plate (for a particular wave-length λ) is that the axis EE′ of the resultant rotation should be on the equator; so that, if 2ω denote the lattitude of E, the faster ellipse, $2\omega = 0$. And if the retardation of the plate is to have a required value δ, the magnitude of the resultant rotation ϕ should be equated to this value.

Both the magnitude ϕ as well as the axis EE′ of the resultant rotation will in general alter with the wave-length, since the magnitudes (though not the axes) of the component rotations are a function of wave-length. Hence, for the combination to be considered achromatic in the immediate neighbourhood of the wave-length λ, some more conditions have to be imposed; when the wave-length is increased by $\Delta\lambda$ the values of the composite phase retardation ϕ, as well as the latitude 2ω and longitude $2l$ of E, should remain unaltered at least to the first order of approximation.

Instead of following matrix methods, we shall use a well-known geometrical construction (illustrated in Fig. 1) for finding the resultant of any

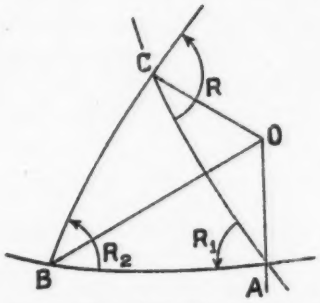

FIG. 1. Composition of Rotations: $2R_1 + 2R_2 = 2R$.

two rotations. If ABC be a spherical triangle described on a sphere whose centre is O then: *a rotation about AO through twice the internal angle at A, followed by a rotation about BO through twice the internal angle at B is equivalent to a rotation about CO through twice the external angle at C.*

3. A BIREFRINGENT COMBINATION WITH VARIABLE RETARDATION

We shall now show that a combination of three birefringent plates, the first and last of which have their corresponding principal planes parallel and their retardations identical, will behave as a single purely birefringent

plate of retardation 2δ, whose fast vibration direction is inclined at an angle c_1 to that of the first plate, where 2δ and c_1 are given by:

$$\cos \delta = \cos 2\delta_1 \cos \delta_2 - \sin 2\delta_1 \sin \delta_2 \cos 2c \qquad (a)$$

and

$$\cot 2c_1 = \operatorname{cosec} 2c \, (\sin 2\delta_1 \cot \delta_2 + \cos 2\delta_1 \cos 2c) \qquad (b)$$

In the above relations $2\delta_1$ is the common retardation of the first and last plates, $2\delta_2$ that of the central plate, and c the angle between the fast vibration directions of the central plate and the other two.

Before proceeding to prove these relations, we may mention that by choosing 2 quarter-wave plates as the first and last plates, and a half-wave plate as the central plate, we can, by rotating the central plate, vary the retardation of the combination continuously from 0 to 2π; the principal planes of the combination will always be inclined at 45° to those of the first plate.

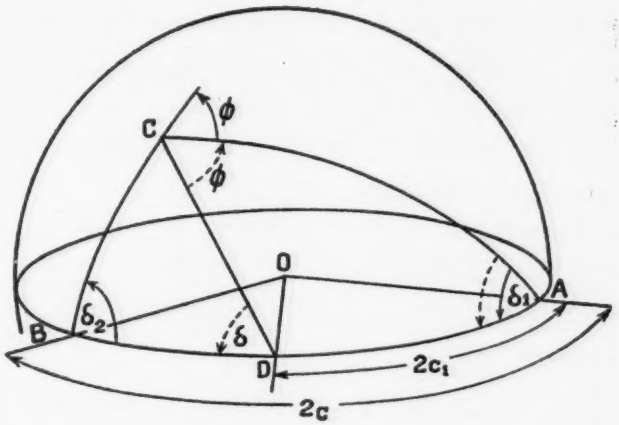

FIG 2.

Referring to Fig. 2, a rotation $2\delta_1$ about AO followed by a rotation $2\delta_2$ about BO, where A and B give the orientation of the fast axes of the first and middle plates respectively, is equivalent to a rotation 2ϕ about CO—where $CAB = \delta_1$, $CBA = \delta_2$ and $\pi - ACB = \phi$. Let us draw an arc CD such that $ACD = \phi$, D being the point of intersection of this arc with the equator. Considering now the triangle ACD, the rotation 2ϕ about C (representing the combined action of the first 2 plates) followed by a rotation $2\delta_1$ about A (which now represents the action of the third plate) is equivalent to a rotation about D through twice the angle external to ADC. The

combination is therefore equivalent to a birefringent plate of retardation 2δ whose fast axis is inclined at c_1 to that of the first plate—where $\delta = \pi - \text{ADC}$, and $2c_1 = \text{arc AD}$.

The expressions which we shall derive for δ and c_1 do not depend on assuming that D lies within AB.

From the triangle ADC, denoting arc CD by d

$$\cos \delta_1 = \cos \phi \cos \delta + \sin \phi \sin \delta \cos d \tag{1}$$

From the triangle DBC

$$\cos \delta_2 = \cos 2\phi \cos \delta + \sin 2\phi \sin \delta \cos d \tag{2}$$

Eliminating d by multiplying (1) by $2 \cos \phi$ and subtracting from (2), we obtain:

$$\cos \delta = 2 \cos \phi \cos \delta_1 - \cos \delta_2$$

The value of $\cos\phi$ to be substituted is obtained from the triangle ABC.

$$\cos \phi = \cos \delta_1 \cos \delta_2 - \sin \delta_1 \sin \delta_2 \cos 2c$$

Hence

$$\cos \delta = \cos 2\delta_1 \cos \delta_2 - \sin 2\delta_1 \sin \delta_2 \cos 2c$$

which is the required expression (*a*).

To get an expression for c_1, we use a well-known trigonometrical theorem, which gives from triangle ABC, on denoting are AC by b

$$\cot 2c \sin b = \cot \text{ACB} \sin \delta_1 + \cos b \cos \delta_1$$

and from triangle ACD

$$\cot 2c_1 \sin b = \cot \text{ACD} \sin \delta_1 + \cos b \cos \delta_1$$

or

$$\sin b (\cot 2c_1 - \cot 2c) = 2 \sin \delta_1 \cot \phi$$

Substituting

$$\sin b = \sin \delta_2 \cdot \frac{\sin 2c}{\sin \phi}$$

we obtain,

$$\sin 2c (\cot 2c_1 - \cot 2c) = \frac{2 \sin \delta_1}{\sin \delta_2} \cos \phi$$

i.e.,

$$\sin 2c \cot 2c_1 - \cos 2c = \frac{2 \sin \delta_1}{\sin \delta_2} (\cos \delta_1 \cos \delta_2 - \sin \delta_1 \sin \delta_2 \cos 2c)$$

or

$$\sin 2c \cot 2c_1 = \sin 2\delta_1 \cot \delta_2 + \cos 2\delta_1 \cos 2c$$

which is the required expression (*b*).

4. Conditions for Achromatism

To make a combination of the type described above achromatic in the immediate neighbourhood of the wave-length λ we may differentiate the expressions on the right-hand side of the equations (*a*) and (*b*), and equate to zero. Since however our object is to attain achromatism over a finite range of wave-lengths we proceed in a different manner.

Let λ' and λ'' be two wave-lengths on either side of λ for which the retardations of each individual plate are multiplied by a factor $(1 - f)$ and $(1 + f)$ respectively. Values corresponding to these two wave-lengths will be denoted by corresponding single- and double-primed symbols.

We impose the conditions that the retardation $2\delta'$ and $2\delta''$ of the combination for these two wave-lengths be equal to one another and to 2Δ, the required retardation of the achromatic combination; also that the orientations of the fast axes of the combination for these two wave-lengths should be the same, *i.e.*, $c_1' = c_1''$. We then get three equations involving the three unknowns δ_1, δ_2 and c:

$$\cos \Delta = \cos 2\delta_1' \cos \delta_2' - \sin 2\delta'_1 \sin \delta_2' \cos 2c \tag{1}$$

$$\cos \Delta = \cos 2\delta_1'' \cos \delta_2'' - \sin 2\delta_1'' \sin \delta_2'' \cos 2c \tag{2}$$

$$\sin 2\delta_1' \cot \delta_2' + \cos 2\delta_1' \cos 2c = \sin 2\delta_1'' \cot \delta_2'' + \cos 2\delta_1'' \cos 2c \tag{3}$$

From these three relations we may get two relations involving only the unknowns δ_1 and δ_2.

Thus eliminating $\cos 2c$ from (1) and (2)

$$\frac{\cos \delta_2' \cos 2\delta_1' - \cos \Delta}{\sin 2\delta_1' \sin \delta_2'} = \frac{\cos \delta_2'' \cos 2\delta_1'' - \cos \Delta}{\sin 2\delta_1'' \sin \delta_2''} \tag{4}$$

For the second equation, we eliminate $\cos 2c$ as well as $\cos \Delta$ between the three equations. Thus equating the right-hand expression of (1) and (2), and then substituting the value of $\cos 2c$ from (3) we get:

$$\frac{\cos 2\delta_1' \cos \delta_2' - \cos 2\delta_1'' \cos \delta_2''}{\sin 2\delta_1' \sin \delta_2' - \sin 2\delta_1'' \sin \delta_2''} = \frac{\sin 2\delta_1' \cot \delta_2' - \sin 2\delta_1'' \cot \delta_2''}{\cos 2\delta_1'' - \cos 2\delta_1'}$$

which on simplification yields,

$$(1 - \cos 2\delta_1' \cos 2\delta_1'')\ (\cos \delta_2' + \cos \delta_2'')$$

$$= \sin 2\delta_1' \sin 2\delta_1''\ \frac{\sin \delta_2' \cos \delta_2'' + \cos \delta_2' \sin \delta_2''}{\sin \delta_2' \sin \delta_2''}$$

Substituting

$$\delta_2' = (1 - f)\ \delta_2 \text{ and } \delta_2'' = (1 + f)\ \delta_2,$$

we get

$$(1 - \cos 2\delta_1' \cos 2\delta_1'')\, 2 \cos \delta_2 \cos f\delta_2 = \frac{\sin 2\delta_1' \sin \delta_1'' \cdot \sin 2\delta_2}{\frac{1}{2}(\cos f\delta_2 - \cos \delta_2)} \qquad (5)$$

We have to find values of δ_1 and δ_2 which simultaneously satisfy (4) and (5). The task is considerably simplified when we note that equation (5) is identically satisfied for $\delta_2 = \pi/2$. The corresponding value of δ_1 may be found from (4), which now becomes:

$$\sin f\frac{\pi}{2} \cdot \sin 2(\delta_1' + \delta_1'') = \cos \Delta (\sin 2\delta_1'' - \sin 2\delta_1')$$

Substituting

$$\delta_1' = (1 - f)\ \delta_1 \text{ and } \delta_1'' = (1 + f)\ \delta_1,$$

this gives

$$\sin f\,.\ 2\delta_1 = \frac{\sin f \cdot \frac{\pi}{2}}{\cos \Delta} \cdot \sin 2\delta_1 \qquad (I)$$

An expression for $\cos 2c$ is got by eliminating Δ from (1) and (2).

$$\cos 2c = \frac{\sin f\frac{\pi}{2}}{\cos f\frac{\pi}{2}} \cdot \frac{\cos 2\delta_1' + \cos 2\delta_1''}{\sin 2\delta_1' - \sin 2\delta_1''}$$

or

$$\cos 2c = -\frac{\tan f\frac{\pi}{2}}{\tan f \cdot 2\delta_1} \qquad (II)$$

For any required retardation 2Δ and for any choice of the arbitrary parameter f which determines the range of achromatism, the retardation $2\delta_1$ of the first and last plates of the achromatic combination may be found from I; since this is a transcendental equation it has to be solved by an iteration procedure. The orientation c of the central plate, which should be a half-wave plate, may then be obtained from II. The orientation c_1' of the principal planes of the combination may, if required, be calculated from

$$\sin 2c \cot 2c_1' = \sin (1 - f)\, 2\delta_1 \tan f\frac{\pi}{2} + \cos (1 - f)\, 2\delta_1 \cos 2c \qquad (III)$$

The retardation and the orientation of the principal planes of the achromatic retardation plate, taken as a whole, will disperse slightly with wavelength. The combination will possess exactly the same retardation 2Δ and the same orientation c_1' of its fast vibration direction, for two wavelengths; these wave-lengths, if we neglect the dispersion of birefringence, will be $(1+f)\lambda$ and $(1-f)\lambda$; and we may, somewhat arbitrarily, consider the range of achromatism as extending from $(1+\sqrt{2}f)\lambda$ to $(1-\sqrt{2}f)\lambda$.

We shall present the numerical solutions only for the case of an achromatic quarter-wave plate—for which we must substitute $2\Delta = 90°$. If we choose $f = \cdot 18$ the solution of I can be shown to be $2\delta_1 = 115° 42'$; so that from II, $\cos 2c = -\cdot 7639$ or $2c = 139° 48'$. The corresponding value of c_1' is got from III, which yields $\cot 2c_1' = \cdot 5489$ or $2c_1' = 61° 14'$.

On the other hand if $f = \cdot 1414$ we get

$$2\delta_1 = 115° 30',\ 2c = 140° 26',\ 2c_1' = 61° 30'$$

Thus an achromatic quarter-wave plate is obtained by superposing three plates of the same birefringent material (the dispersion of the birefringence of which need not be negligible). The central plate should be a half-wave plate for the wave-length λ in the middle of the spectral range to be covered. The first and last plates should have their principal planes in parallel orientation and their retardations of the same magnitude—equal to, say, $2\delta_1$ for the wave-length λ. With respect to the common fast vibration direction of the first and last plates, let the inclinations of the fast vibration direction of the central plate, and of the combination taken as a whole, be c and c_1' respectively. Then for covering the range of wave-lengths from about $1\cdot 25\,\lambda$ to $\cdot 75\,\lambda$ the set of values, $2\delta_1 = 115° 42'$, $c = 69° 54'$ may be used—for which $c_1' = 30° 37'$. The achromatism is not high but can be increased by restricting the range to be covered. Thus for covering the region from $1\cdot 2\,\lambda$ to $\cdot 8\,\lambda$ the values suitable are:

$$2\delta_1 = 115° 30',\ c = 70° 13',$$

for which

$$c_1' = 30° 45'$$

We may mention that the range over which these combinations may be considered achromatic would have to be altered if the dispersion of the birefringence is appreciable, but their essentially achromatic nature would not be changed. For retardation plates of mica, the dispersion of birefringence is negligible.[2]

EXPERIMENTAL VERIFICATION

Two plates of mica were cleaved, whose thicknesses were in the ratio of 180 : 115·75, the former being nearly a half-wave plate for the D lines of sodium. The second plate was cut into two portions which were then superposed and cemented together at a corner—the principal planes of the two portions being made exactly coincident while the cement was still wet. After the cement was dry, the half-wave plate—with a drop of copal varnish on either side—was inserted *between* the other two plates at roughly the calculated angle. The combination was placed on a mirror and, with the aid of another mirror, white light from a point source was made to pass normally through a polaroid and the combination of plates and then back again—after which it reached the eye. The orientation of the central plate was then gradually altered while the varnish was still wet—and a position was found where the image of the point source could be almost completely extinguished by rotating the polaroid. It was also verified that the combination thus prepared showed two practically black extinction positions when rotated between crossed polaroids.

The author's thanks are due to Prof. Sir C. V. Raman for his kind interest in this work.

SUMMARY

An achromatic quarter-wave plate is obtained by superposing three birefringent plates of the same material; the first and last should have the same retardation $2\delta_1$, their fast vibration directions being parallel to one another but inclined at a specific angle c to that of the central plate—of retardation π. The desired range of achromatism determines the optimum values of $2\delta_1$ and c (which, in turn, will determine the orientation of the effective principal planes of the combination). As an example, using mica retardation plates prepared for Hg 5461, the range from 4100 Å to 6800 Å is covered with $2\delta_1 = 115° 42'$ and $c = 69° 54'$.

Further, for a particular wave-length, a birefringent compensator of variable retardation (0 to 2π) is obtained by interposing a half-wave plate that can be rotated in its own plane, between two quarter-wave plates that have their fast vibration directions parallel.

The results follow from the Poincaré sphere by geometrically compounding successive rotations.

REFERENCES

1. Pockels .. *Lehrbuch der Kristalloptik, Teubner,* 1906, 280.
2. Mathieu .. *Bull. Soc. Franc. Min.,* 1934, **57**, 233.

INVERSE PROBABILITY AND EVOLUTIONARY MARKOFF STOCHASTIC PROCESSES*

By Alladi Ramakrishnan

(*Department of Physics, University of Madras*)

Received February 19, 1955

Introduction

The subject of inverse probability has unfortunately been regarded with general suspicion and so not many attempts have been made to apply the theory even to cases where unambiguous answers can be given, on a proper mathematical formulation of the problems. It is the object of this paper to emphasise that in the case of evolutionary stochastic processes, the application of Bayes' theorem yields results as meaningful and rigorous as could be obtained by the application of the concepts of ordinary (as contrasted with inverse) probability theory.

We shall first formulate the basic problem of inverse probability in relation to univariate stochastic processes and point out why it was erroneously supposed to be a difficult problem. We shall then show how the problem can be solved rigorously by reconsidering the Chapman-Kolmogoroff equation and consequently the stochastic differential equations derived therefrom, in a new light.

2. Formulation of the Problem

Let us consider a stochastic process representing the 'evolution' in a Markovian manner with respect to the one-dimensional parameter t of the probability distribution of a continuous stochastic variate $x(t)$. We are thus interested in the probability frequency function (p.f.f.) $\pi(x_1|\, x_0;\ t_1, t_0)\, dx_1$ representing the probability that $x(t_1)$ lies between x_1 and $x_1 + dx_1$ given that $x(t_0) = x_0$, $(t_1 > t_0)$.

It shall be assumed that the process is homogeneous in t and so π is a function only of the interval $t_1 - t_0$ and not t_1, t_0 severally. Though it is customary to set $t_0 = 0$ and write $t_1 - t_0$ as t_1, we shall deliberately refrain from doing so, for reasons which will be clear presently. On the contrary we shall use the notation $\pi(x;\ t)$ to represent the p.f.f. of $x\ (t)$ when we do

* Read at the Annual Meeting of the Indian Academy of Sciences at Belgaum, December, 1954.

not specify the "initial" conditions, *i.e.*, the p.f.f. at some t', $(t' < t)$†. Thus by definition we immediately write

$$\pi(x_1;\ t_1) = \int_{x_0} \pi(x_0;\ t_0)\ \pi(x_1|\ x_0;\ t_1, t_0)\, dx_0,\ (t_0 < t_1). \tag{1}$$

When we deal with distributions at two different values of t, say t_1 and t_0, $(t_1 > t_0)$, we shall call the distribution at t_1 the emergent spectrum with respect to the injected spectrum at t_0, and $\pi(x_1|\ x_0;\ t_1, t_0)$ the fundamental transition probability for the finite interval $t_1 - t_0$. It is quite clear that $\pi(x_1|\ x_0;\ t_1, t_0)$ is only a particular case of $\pi(x_1;\ t_1)$ where the injection spectrum at t_0 is given by $\pi(x;\ t_0) = \delta(x - x_0)$ where δ is the Dirac delta function. π satisfies the Chapman-Kolmogoroff equation,

$$\pi(x;\ t_1) = \int_{x'} \pi(x';\ t_0 + \tau)\ \pi(x|\ x';\ t_1,\ t_0 + \tau)\, dx' \tag{2}$$

for all values of τ such that $t_0 < t_0 + \tau - t_1$ and in particular

$$\pi(x_1|\ x_0;\ t_1,\ t_0) = \int_{x'} \pi(x'|\ x_0; t_0 + \tau,\ t_0)\ \pi(x_1|\ x';\ t_1, t_0 + \tau)\, dx'. \tag{3}$$

By making $t_0 + \tau \to t_1$, we obtain the well-known forward differential equation of Kolmogoroff for π provided we know the limit of $\pi(x_1|\ x';\ t_1, t)$ as $t_1 - t \to +0$, *i.e.*, to speak in physical terms we know what 'happens' to the stochastic variable is an infinitesimal interval Δ.

From equation (1) we find that the emergent p.f.f. $\pi(x_1;\ t_1)$ can be computed if we know the injected p.f.f. $\pi(x_0;\ t_0)$ and the fundamental transition probability $\pi(x_1|\ x_0;\ t_1, t_0)$. But the computation of $\pi(x_0;\ t_0)$ given $\pi(x_1;\ t_1)$ seems at first a formidable task since it involves the solution of equation (1) which is a Volterra integral equation of the first kind. It is the principal object of the present paper to show that no complex procedure is necessary to solve for the injected spectrum *in view of the peculiar properties of the fundamental transition probability*. We shall prove by very simple arguments that the inversion of (1) is given by

$$\pi(x_0;\ t_0) = \int_{x_1} \pi(x_1;\ t_1)\ \hat{\pi}(x_0|\ x_1;\ t_1,\ t_0)\, dx_1,\quad (t_1 > t_0), \tag{4}$$

where $\hat{\pi}(x_0 \mid x_1;\ t_1,\ t_0)$ is a function obtained by replacing $t_1 - t_0$ *by the negative quantity* $t_0 - t_1$, in $\pi(x_0|\ x_1;\ t_0,\ t_1)$ which we remember is only a function of $t_1 - t_0$ and not t_1, t_0 severally. $\hat{\pi}$ *unlike* π *has no probability significance and is only a functional operator.*

† Throughout this paper π denotes a p.f.f. Distinction between different distributions will be obvious from the variable used to define π.

In equation (1) since $\pi(x;\ t)$ and $\pi(x_1|\ x_0;\ t_1, t_0)$ are both probability densities, they are positive and so $\pi(x;\ t_1)$ exists for any $t_1 > t_0$, *i.e.*, $t_1 - t_0$ is positive and finite and can be chosen as large as we please. If $\pi(x|\ x_0;\ t_1,\ t_0)$ tends to a stationary distribution as $t_1 - t_0 \to +\infty$, so also does $\pi(x_1;\ t_1)$.

On the contrary in (4) since $\hat{\pi}$ is not a p.f.f. $\pi(x;\ t_0)$ does not remain a p.f.f. for all values of $t_0 < t_1$. It shall be proved that there exists a t_p such that $\pi(x_0;\ t_0)$ is a p.f.f. in the domain $t_p \leqslant t_0 \leqslant t_1$ and the process cannot be traced back to a point $t < t_p$.

3. Solution of the Problem

It is necessary to recall the method of derivation of the forward differential equation for $\pi(x;\ t)$ from the Chapman-Kolmogoroff equation. In particular, we first consider the simple case when

$$\pi(x_1|\ x_0;\ t_1, t_0) \to \mathrm{R}(x_1|\ x_0)\,\Delta + \delta(x_1 - x_0)\,(1 - \Delta \int_{x'} \mathrm{R}(x'|\ x_0)\,dx') \quad (5)$$

as $t_1 - t_0 \to \Delta \to +0$.

In such a case, (1) assumes the form

$$\pi(x;\ t_0 + \Delta) = \Delta \int_{x_0} \pi(x_0;\ t_0)\ \mathrm{R}(x|\ x_0)\,dx_0 + \pi(x;\ t_0)\,\{1 - \Delta \int_{x'} \mathrm{R}(x'|\ x)\,dx'\} + 0(\Delta^2). \quad (6)$$

(6) can be reduced to a differential equation in t by making $\Delta \to +0$. From (6) $\pi(x;\ t_0 + \Delta)$ can be computed if $\pi(x;\ t_0)$ is known, and thus the emergent p.f.f. $\pi(x;\ t_1)$ is obtained from $\pi(x_0;\ t_0)$ by a process of integration, *i.e.*, by solving the differential equation in t.

We notice we can write (6) in the following form:

$$\pi(x;\ t_0 + \Delta) = \Delta \int_{x_0} \pi(x_0;\ t_0 + \Delta)\ \mathrm{R}(x|x_0)\ dx_0 + \pi(x;\ t_0) - \pi(x;\ t_0 + \Delta) \cdot \Delta \int_{x'} \mathrm{R}(x'|\ x)\,dx' + 0(\Delta^2) \quad (7)$$

which can be re-written in the equivalent form (8) by using the device of writing $t_0 - \Delta$ and t_0 for t_0 and $t_0 + \Delta$ respectively.

$$\pi(x;\ t_0 - \Delta) = \pi(x;\ t_0) + \Delta\,\pi(x;\ t_0) \int_{x'} \mathrm{R}(x'|\ x)\,dx' - \Delta \int_{x_0} \pi(x_0;\ t_0)\ \mathrm{R}(x|x_0)\,dx_0. \quad (8)$$

In this form we immediately realise $\pi(x;\ t_0 - \Delta)$ *can be obtained from* $\pi(x;\ t_0)$ *by merely replacing* Δ *by* $-\Delta$ *in* $\pi(x_0;\ t + \Delta)$. Thus $\pi(x_0;\ t_0)$ can be obtained by a process of integration from $\pi(x_1;\ t_1)$, $(t_1 > t_0)$.

Considering processes more general than that defined by (1), Bartlett (see Ramakrishnan, 1952) has observed that the forward differential equation for π can be written in the *generalised symbolic form*,

$$\frac{\partial \pi(t_0 + \tau)}{\partial \tau} = \mathrm{F}\{\pi(t_0 + \tau)\}, \tag{9}$$

(τ positive) omitting for the moment the argument x to indicate that the right-hand side may involve several values of the argument. If $\pi(t_0)$ is a p.f.f. $\pi(t_0 + \tau)$ is a p.f.f. for all τ from 0 to $+\infty$. Noting that the above equation is derived from the following

$$\pi(t_0 + \tau + \Delta) = \pi(t_0 + \tau) + \mathrm{F}\{\pi(t_0 + \tau)\} \cdot \Delta + 0(\Delta^2), \tag{10}$$

we re-write it in the form

$$\pi(t_0 + \tau + \Delta) = \pi(t_0 + \tau) + \mathrm{F}\{\pi(t_0 + \tau + \Delta)\} \cdot \Delta + 0(\Delta^2). \tag{11}$$

Therefore using the same device as before of writing $t_0 - \overline{\tau + \Delta}$ and $t_0 - \tau$ in the place of $t_0 + \tau$ and $t_0 + \tau + \Delta$,

$$\pi(t_0 - \overline{\tau + \Delta}) - \pi(t_0 - \tau) = -\Delta\, \mathrm{F}\{\pi(t_0 - \tau)\}$$

or

$$\frac{\partial \pi(t_0 - \tau)}{\partial \tau} = -\mathrm{F}\{\pi(t_0 - \tau)\}. \tag{12}$$

In other words $\pi(t_0 - \tau)$ is obtained from $\pi(t_0)$ by merely replacing τ by $-\tau$ in $\pi(t_0 + \tau)$.

It now remains to show that this process of inversion, *i.e.*, determining the p.f.f. at $t_0 - \tau$ from the p.f.f. at t_0 cannot be carried on for all values of τ except in the trivial case when the emergent distribution at t_0 is a stationary distribution.

To do this we first consider the simpler case defined by (5) and examine the essential difference between (6) and (7). Noting that $\mathrm{R}(x|\, x_0)$ and $\pi(x;\ t_0)$ are non-negative in (6), we find that $\pi(x;\ t_0 + \Delta)$ is also non-negative. On the contrary $\pi(x;\ t_0 - \Delta)$ will be negative for such values of x for which $\pi(x;\ t_0) = 0$ if $\mathrm{R}(x|\, x')$ is positive for some value of x'. In such a case $\pi(x;\ t_0)$ cannot be an ' emergent ' distribution, *i.e.*, it cannot be traced back. *Thus there exists a smallest value* τ_p *for* τ *such that* $\pi(x;\ t_0 - \tau_p)$

is zero for some value of x. Thus $t_p = t_0 - \tau_p$ is the earliest point to which we can trace back the process. Identical arguments apply to a π function satisfying Bartlett's generalised symbolic equation. We call t_p the '*primitive origin*' of the process and τ_p the '*age*' of the process at t_0. The distribution at any point $t_0 - \tau$, $0 < \tau < \tau_p$ is given by equation (12), *i.e.*, if we inject $\pi(x;\ t_0 - \tau)$ satisfying (12), then it will emerge as $\pi(x;\ t_0)$ at t_0. It also follows that $\pi(x;\ t_0 - \tau_1)$, the injection spectrum at $t_0 - \tau_1$, is the emergent spectrum with respect to $\pi(x;\ t_0 - \tau_2)$ the corresponding injection spectrum at $t_0 - \tau_2$, $(\tau_2 > \tau_1)$. $\pi(x;\ t - \tau)$ satisfies the modified Chapman-Kolmogoroff equation

$$\pi(x;\ t_0 - \tau) = \int_{x'} \pi(x';\ t_0 - \tau')\ \pi(x|\ x';\ t_0 - \tau,\ t_0 - \tau')\ dx', \quad (13)$$

where

$$t_0 - \tau_p < t_0 - \tau' < t_0 - \tau$$

i.e.,

$$\tau_p > \tau' > \tau.$$

The distribution at $t_0 - \tau_p$ can only be an injected distribution and cannot be the emergent distribution of a process started 'earlier'. Therefore we call it the primitive *a priori* spectrum.

Illustration in the case when $x(t)$ *is discrete.*—We shall now consider the case when $x(t)$ is discrete and can assume mutually exclusive values $x_1, x_2, \ldots, x_j, \ldots$ and define $\pi(j;\ t)$ as the probability (not probability density) that $x(t) = x_j$. If $\pi(j|\ i;\ t_1,\ t_0) \to \mathrm{R}(j|\ i) \cdot \Delta$ as $t_1 - t_0 \to \Delta \to +0$ it is well known that π satisfies the matrix differential equation [*cf.* integro-differential equation when $x(t)$ is continuous]

$$\frac{\partial \vec{\pi}(j;\ t)}{\partial t} = [\mathrm{R}]\,\vec{\pi}(j;\ t), \quad (14)$$

where $\vec{\pi}(j;\ t)$ is a column vector with components corresponding to different values of j and [R] is the matrix with non-diagonal elements $\mathrm{R}(j|\ i)$ and the diagonal elements defined by $\mathrm{R}(i|\ i) = -\sum_j \mathrm{R}(j|\ i)$. The solution of the above equation is given by

$$\vec{\pi}(j;\ t_1) = e^{[\mathrm{R}](t_1 - t_0)}\,\vec{\pi}(j;\ t_0). \quad (15)$$

By considering the infinitesimal transformation $\pi(j;\ t)$ to $\pi(j;\ t + \Delta)$ we note that if, $\pi(j;\ t) \geqslant 0$ for all j so is $\pi(j;\ t + \Delta)$. Thus $\pi(j;\ t_1) \geqslant 0$ for all j for all $t_1 > t_0$ provided $\pi(j;\ t_0) \geqslant 0$ for all j.

By arguments now familiar from the previous section $\pi(j;\ t-\Delta)$ can be obtained from $\pi(j;\ t)$ and so

$$\overrightarrow{\pi}(j;\ t_0)=e^{-[R](t_1-t_0)}\overrightarrow{\pi}(j;\ t_1), \qquad (t_1>t_0) \tag{16}$$

i.e., given the emergent distribution $\overrightarrow{\pi}(j;\ t_1)$ we compute the injected distribution at t_0. This computation can be done for t_0 in the domain $t_1-\tau_p \leqslant t_0 \leqslant t_1$. At $t_0=t_1-\tau_p$, $\pi(j;\ t_1-\tau_p)=0$ at least for one value of j.

To illustrate this we shall take a system capable of assuming one of four states, *i.e.*, $j=1, 2, 3, 4$. We shall take the following values for R:

$$\begin{pmatrix} R(1|1) & R(1|2) & R(1|3) & R(1|4) \\ R(2|1) & R(2|2) & R(2|3) & R(2|4) \\ R(3|1) & R(3|2) & R(3|3) & R(3|4) \\ R(4|1) & R(4|2) & R(4|3) & R(4|4) \end{pmatrix} = \begin{pmatrix} -12 & 0 & 0 & 0 \\ 1 & -8 & 0 & 0 \\ 7 & 6 & -4 & 0 \\ 4 & 2 & 4 & 0 \end{pmatrix}$$

We assume that at some t_1, the values of $\pi(j;\ t_1)$ are as given below: then the primitive origin $t_1-\tau_p$ is determined.

$$\left.\begin{aligned} &\pi(1;\ t_1)=\cdot 1,\ \pi(1;\ t_1-\tau_p)=\cdot 3729 \\ &\pi(2;\ t_1)=\cdot 1,\ \pi(2;\ t_1-\tau_p)=\cdot 2074 \\ &\pi(3;\ t_1)=\cdot 2,\ \pi(3;\ t_1-\tau_p)=0 \\ &\pi(4;\ t_1)=\cdot 6,\ \pi(4;\ t_1-\tau_p)=\cdot 4197 \end{aligned}\right\} \text{ for } \tau_p=\cdot 1097$$

Conditional Inverse Probabilities.—We have observed that the p.f.f. $\pi(x_1|\,x_0;\ t_1,\ t_0)$, $(t_1>t_0)$ is a particular case of $\pi(x_1;\ t_1)$ with $\pi(x;\ t_0)=\delta(x-x_0)$. In this case the process cannot be traced back further than t_0, *i.e.*, to a point $t<t_0$, since $\pi(x;\ t_0)=0$ for $x\neq x_0$. Hence $\pi'(x;\ t_0-\Delta)$ is negative (Δ positive) for $x\neq x_0$. Therefore a delta function can only be an injected spectrum and cannot be the emergent spectrum of a stochastic process. Thus $\hat{\pi}(x|\,x_0;\ t,\ t_0)$, $(t>t_0)$ has no probability interpretation.

On the other hand, in a previous paper, the author (1954) defined P $(x_0|\,x_1;\ t_0,\ t_1)$ as the conditional inverse probability that $x(t_0)$ lies between x_0 and x_0+dx_0 given that $x(t_1)=x_1$, $(t_0<t_1)$. This has to be distinguished from $\hat{\pi}$. *When in defining P we state that $x(t_1)=x_1$, we mean that the observed value of the emergent spectrum at t_1 is x_1 and not that the emergent spectrum is a delta function.* In $\hat{\pi}$ we require the emergent spectrum to be a delta function and this is not possible. Hence $\hat{\pi}$ is not a probability magnitude.

In his previous contribution, the author showed that $P(x_0|\, x_1;\ t_0,\ t_1)$ cannot be obtained unless the injected spectrum at a certain $t < t_0$ is known. If that is known to be $\pi(x;\ t)$ then by considering the joint probability of obtaining $x(t) = x$, $x(t_0) = x_0$ and $x(t_1) = x_1$ and applying Bayes' theorem, it was shown that

$$P(x_0|\, x_1;\ t_0,\ t_1) = \frac{\int \pi(x;\ t)\, \pi(x_0|x;\ t_0,\ t)\, \pi(x_1|x_0;\ t_1,\ t_0)\ dx}{\int \pi(x;\ t)\ \pi(x_1|\, x;\ t_1,\ t)\, dx}$$

$$= \frac{\pi(x_0;\ t_0)\ \pi(x_1|\, x_0;\ t_1,\ t_0)}{\pi(x_1;\ t_1)}, \quad (t < t_0 < t_1) \qquad (17)$$

If $\pi(x;\ t)$ is known $\pi(x_0;\ t_0)$ and $\pi(x_1;\ t_1)$ can be computed from the fundamental transition probabilities corresponding to the intervals $t_0 - t$ and $t_1 - t_0$. Hence the method of computation of P described in that paper is equivalent to the present one.

In this paper we have shown that if $\pi(x;\ t)$ is known at some point, it can be extended to the entire permissible domain of t and so P can be computed if the distribution at some t, and of course the fundamental transition probability for a finite interval are known.

Problems Involving a Probability Frequency Function for the Primitive Origin

Till now we have assumed that the distribution at t is emergent with respect to that at $t - \Delta$ (except when t is the primitive origin) and injected with respect to that at $t + \Delta$. Or, in other words, given the distribution at t_0, the injected distribution at $t < t_0$ is a function of $t_0 - t$. The primitive origin therefore is a definite point and is uniquely determined if we know the probability frequency function at t_0.

We shall now consider the following class of problems. We are given the probability frequency function at t_0. The primitive origin of the process lies between $t_0 - \tau$ and $t_0 - (\tau + d\tau)$ with probability $\eta(\tau)\, d\tau$. At this origin a p.f.f. $\phi(x)$ independent of τ is injected. Given $\eta(\tau)$ our problem is to determine the injection spectrum $\phi(x)$.

Using simple probability arguments we write

$$\pi(x;\ t_0) = \int_0^\infty d\tau \int_{x'} \phi(x')\, \pi(x|\, x';\ t_0,\ \tau)\, \eta(\tau)\ dx' \qquad (18)$$

Since $\pi(x;\ t_0)$ and $\pi(x|\, x',\ t_0,\ \tau)$ are known, $\eta(\tau)$ or $\phi(x')$ can be determined if the other is known, by the classical theory of integral equations;

but the explicit solution of $\phi(x')$ or $\eta(t)$ amenable to numerical computation is a very difficult problem. On the contrary, the computation of π given ϕ is a much easier problem. Considering the very simple special case when x is a discrete variable representing a Poisson distribution

$$\pi(x|\, x';\ t,\ t-\tau) = \frac{e^{-\lambda\tau}(\lambda\tau)^{x-x'}}{(x-x')!}$$

and is independent of t,

$$\pi(x;\ t_0) = \int_0^\infty e^{-\lambda\tau}\ \frac{(\lambda\tau)^x}{x!}\ \eta(\tau)\ d\tau \ddagger \tag{19}$$

As expected $\pi(x;\ t_0)$ is independent of t_0 since $\pi(x|\, x';\ t,\ t-\tau)$ is only a function of τ. The computation of η given π is still a difficult problem.

Consael has discussed a class of distributions defined by

$$\pi(x;\ t) = \int_0^\infty e^{-\lambda t}\ \frac{(\lambda t)^x}{x!}\ \eta(\lambda)\ d\lambda \tag{20}$$

The physical meaning of this class of distributions is obvious. The parameter λ of the Poisson distribution corresponding to an interval t is itself a stochastic variable with a p.f.f. $\eta(\lambda)$. Comparing (20) with (19), we find that all the results of Consael apply to processes defined by (19) provided we replace λ by t, since the Poisson distribution is symmetrical with respect to t and λ.

Problems involving a "switch-off" of the process.—Let t_0 be the primitive origin of a process where we inject the primitive p.f.f. $\phi(x)$. Let us assume that we "switch-off" the process between $t_0+\tau$ and $t_0+\tau+d\tau$ with probability $\eta(\tau)\,d\tau$. After "switch-off" the stochastic variable retains the value it had at τ till infinity. Thus the p.f.f. at $t_0+\tau$ as $\tau\to\infty$ is given by

$$\pi(x';\ \infty) = \int_a dx \int_0^\infty \phi(x)\,\pi(x'|\, x;\ t_0+\tau,\ t_0)\,\eta(\tau)\,d\tau \tag{21}$$

Comparing this equation with (18) we find $\pi(x';\ \infty)$ is identical with $\pi(x;\ t_0)$ of the previous problem. It is interesting to note that in the previous case the point of origin has a p.f.f. while here the point of switch-off has a p.f.f. and therefore the duration of the process has the same p.f.f. in both cases !

Incidentally the author wishes to mention that the process of acceleration of cosmic ray particles in the galactic magnetic field proposed by Fermi is a process which can be interpreted to belong to either of the two equivalent types discussed above.

‡ We assume, without loss of generality, that $x'=0$.

Process involving "absorption".—Finally, we advert to an interesting problem in inverse probability occurring in physics (relating to the range of a fast particle passing through matter) which for some unknown reason has not been investigated in a rigorous manner.

A fast particle of initial energy E_0 passes through matter and loses energy through ionisation or radiation. When it reaches an energy zero or some critical value E_c we say the particle is 'absorbed' or stopped, and the distance t of penetration (considering the passage to be one-dimensional) till it is absorbed is called its range. The author has, in a separate contribution to this journal, derived an expression for $P(E_0; t)\,dt$, the probability that a particle of initial energy E_0 is absorbed between t and $t + dt$ if we know the fundamental transition probability $\pi(E|E_0; t, 0)$.

If we now assume that the initial energy of the particle is a stochastic variable defined by the p.f.f. $\phi(E_0)$, *i.e.*, if $\phi(E_0)$ is the primitive injected spectrum, the p.f.f. of the range is given by

$$P(t)\,dt = \int_{E_0} \phi(E_0)\ P(E_0; t)\ dE_0 \qquad (22)$$

Given $P(t)$ and $P(E_0; t)$ it is theoretically possible to compute the primitive injected spectrum $\phi(E_0)$ but this is not as easy as the computation of $\phi(E_0)$ if we are given the emergent spectrum at t.

There still remains the further question of conditional inverse probability. Given that the observed range was t_0 to compute the probability $\phi_{t_0}(E_0)\,dE_0$ that the initial energy of the particle lies between E_0 and $E_0 + dE_0$. By the simple application of Bayes' theorem we find $\phi_{t_0}(E_0)$ is given by

$$\phi_{t_0}(E_0) = \frac{\phi(E_0)\,P(E_0; t)}{P(t)} = \frac{\phi(E_0)\,P(E_0; t)}{\int \phi(E_0)\,P(E_0; t)\,dE_0} \qquad (23)$$

It is clear that the conditional inverse probability cannot be obtained unless we know $P(t)$ and $P(E_0; t)$ or $\phi(E_0)$ and $P(E_0; t)$. In the opinion of the author, the function $\phi_{t_0}(E_0)$ is physically more important than $P(E_0; t)$ since in many experiments the energy of the particle is estimated from the range and somewhow this aspect seems to have attracted little attention till now.

References

Alladi Ramakrishnan	..	*Proc. Camb. Phil. Soc.*, 1952, **48**, 451.
Consael, R.	..	*Bulletin de la class. des Sciences; Acad. Roy. de Belg.*, 1952, **38**, 442.
Alladi Ramakrishnan and Mathews, P. M.		*Proc. Ind. Acad. Sci.*, 1953, **38A**, 450.
Alladi Ramakrishnan	,,	*Ap. J.*, 1954, **119**, 443.

SCATTERING MEASUREMENTS IN NUCLEAR EMULSIONS AND THEIR APPLICATION TO MEASURING THE CHARGE OF PRIMARY COSMIC RAY NUCLEI

By S. Biswas, B. Peters, F.A.Sc. and Rama

(*Tata Institute of Fundamental Research, Bombay-1*)

Received April 6, 1955

Abstract

Small local dislocations in processed emulsions give rise to spurious scattering whose effect on tracks is indistinguishable from the multiple scattering produced by coulomb interaction. Because of this effect the atomic number of fast primary cosmic ray nuclei will be underestimated in more than 50% of the cases in experiments where multiple scattering of tracks and ionization measurements are employed for charge determination. The spurious scattering was investigated by measuring the track contours of 100 very long tracks due to energetic primary particles; the effect was present in comparable strength in all plates and all types of emulsions which have been investigated. It is too small to affect measurements on medium energy particle tracks (protons with energy below 600 MeV and heavy nuclei with energy below 300 MeV/nucleon). It dominates, however, other sources or error and noise for tracks of energetic particles although it does not preclude occasional observations of very low scattering values. New methods for measuring various forms of noise have also been developed in the course of this work and the noise level for scattering measurements has been reduced below previously accepted values.

Spurious scattering is presumably largely responsible for the discrepancies which appear when one compares the primary charge and energy spectra derived from experiments involving scattering measurements with the corresponding spectra derived from numerous other experiments which employ different techniques.

If the experiments based on scattering measurements are omitted, the remaining evidence strongly favours a spectrum in which the energy per nucleon is nearly independent of atomic number for all primaries. It also favours a charge distribution which has a pronounced minimum for charges $3 \leqslant Z \leqslant 5$ and, therefore, yields a fairly low upper limit for the amount of interstellar matter traversed by primary cosmic ray nuclei.

1. Introduction

The determination of multiple coulomb scattering of particle tracks in nuclear emulsion has been widely used in conjunction with ionization or

range measurements for obtaining information on mass, charge and energy of particles in cosmic ray experiments. The conclusions derived from such scattering measurements can often be confirmed by independent evidence from other sources. At other times, however, conclusions derived from scattering measurements differ from those obtained by other techniques. Such a discrepancy arose for instance, in experiments designed to determine the charge and energy distribution of heavy nuclei in the primary cosmic radiation.

Experiments not involving scattering measurements have been carried out with nuclear emulsions,[1–4] ionization chambers,[5] proportional counters,[6] cloud chambers triggered by proportional counters,[7] and Czerenkov detectors incorporated into an ordinary counter telescope[8]; they support or are consistent with the conclusion that all components of the primary cosmic radiation have closely similar energy spectra (when expressed in terms of energy per nucleon) and that next to protons and α-particles the group of elements comprising Carbon, Nitrogen and Oxygen is strongest while Lithium, Beryllium and Boron nuclei are either rare or absent.

On the other hand experiments involving measurements of multiple scattering in nuclear emulsions lead to an energy spectrum in which the heavy primary component falls off more rapidly with energy than the total primary radiation and yield a charge spectrum in which Li, Be, B nuclei are approximately as abundant as C, N, O nuclei.[9–12]

This discrepancy has led us to re-examine the limitations of the existing techniques for the measurements of multiple coulomb scattering in nuclear emulsions. We found that the discrepancy is apparently due to an imperfection in the nuclear emulsions now in use. This result seems at first sight surprising in view of the fact that numerous experiments have established the reliability of conclusions drawn from scattering measurements on tracks in nuclear emulsions. In particular it has been shown that the mass of slow particles can be correctly derived from the multiple coulomb scattering of their tracks in conjunction with measurements of range or specific ionization. It has also been shown that scattering measurements give the correct answer when applied to tracks of fairly fast particles of known mass and energy, accelerated in the laboratory. Finally, it has been shown that occasionally tracks of particles which are believed to initiate high energy nuclear interactions exhibit very small multiple coulomb scattering, so small that one can set the lower limit for the energy of these particles at 10 BeV or even higher.

However, the imperfection which we have found to be the property of all emulsions so far investigated and which gives rise to *Spurious Scattering*,

does not invalidate any of these well established conclusions. The spurious scattering is too small to interfere with measurements on slow particles and sometimes its effect becomes small for very flat tracks. It does, however, play an important role when measuring tracks of particles whose energy exceeds 0·6 BeV ($p\beta = 1$ Bev/c) in case of protons and 0·3 BeV/nucleon ($p\beta = 0{\cdot}5$ BeV/c) in case of heavy ions, especially if these particles enter the emulsion with dip angles larger than 2°.

We have found this spurious scattering effect present in all emulsions so far investigated: in glass backed emulsions, varying in thickness from 250 μ to 1,000 μ, and in stripped emulsions belonging to various batches. We have also found it in plates developed in various laboratories (Bombay, Bristol, Chicago and Melbourne) and in emulsions with measurable distortion, as well as in others in which distortion was extremely small. There seems to be no correlation between general distortion and spurious scattering. So far we have restricted our investigations to Ilford G-5 emulsions. We have not yet come across plates which are free from the spurious scattering effect or in which the effect is substantially smaller than average.

The small-scale dislocations which give rise to spurious scattering are illustrated in Fig. 1. This figure shows the contours of two α-particle tracks which run closely parallel and are never separated by more than 40 μ. (The ordinate scale in Fig. 1 has been chosen so as to amplify the y-co-ordinate 4,500 times.) The two α-particles belong to a shower containing six relativistic α-particles which survived the destruction by collision of an energetic heavy primary nucleus. The contour of each track was measured in 250 μ steps, but the tracks were measured on different parts of the microscope stage motion. The very close correspondence between these two complicated contours can, therefore, not be attributed to irregularities in the screw which moves the stage. It can obviously not be attributed to causes which affect each track independently like reading errors, grain noise or coulomb scattering of the particles. It is an illustration of the extremely complicated small-scale dislocations which occur in emulsions, quite different from the C or S-shaped distortions which one meets occasionally. The effect illustrated in Fig. 1 exists in all emulsions and for all tracks which we have investigated so far, although in the example which we have chosen for illustration it is somewhat more pronounced than average. Still most of the dislocations common to both the tracks are not larger than fractions of microns, and the overall S or C-shaped distortion is moderate (50 covans).

It is the purpose of this paper to study the effect of these dislocations and of the resulting spurious scattering on experiments designed to measure

multiple coulomb scattering of particles. This is done in Section III. As far as we know, this effect has not previously been discussed and has not been considered in the numerous papers which deal with the multiple scattering technique in nuclear emulsions.

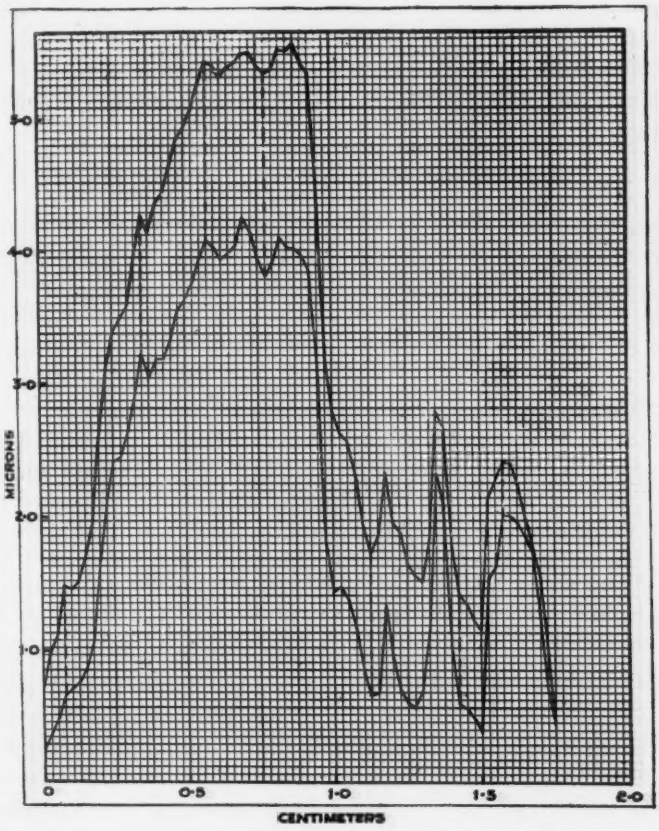

FIG. 1. The contours of two parallel tracks separated by $\sim 40\mu$ were measured at 250μ intervals on different parts of the microscope stage. The tracks belong to two (out of six) relativistic α-particles and are fragments belonging to a heavy nucleus with 15–20 BeV/nucleon destroyed in a nuclear collision. Scattering data are given in Table IV. Correlation between the two contours is due to the local dislocations which give rise to spurious scattering. The ordinate in Fig. 1 has been amplified 4,500 times.

In order to study the effect of spurious scattering we have discussed in Section II, various other limitations on measuring accuracy which arise in the form of "noise" and are fairly well known to workers in this field.

We believe, however, that we have made some improvements in the methods of measuring these various forms of noise and achieved an appreciable reduction in noise level below that previously reported in the literature.

In the last section (Section IV) we consider the effect of spurious scattering on the group of experiments designed to measure the energy and charge distribution of heavy primary cosmic rays. We briefly review the existing evidence and recalculate the upper limit of interstellar material traversed by the cosmic ray beam in the light of present knowledge of its chemical composition.

II. NOISE CORRECTION IN SCATTERING MEASUREMENTS

Throughout this work we have used the co-ordinate method of scattering developed by Fowler.[13] A track is divided into equal cells and measurements consist in determining the distances y_n of the dividing points from a straight line which is reasonably parallel to the track. The mean absolute value of the second co-ordinate difference $\bar{D}$ is defined by:

$$\bar{D} = < | y_n - 2y_{n+1} + y_{n+2} | > \qquad (1)$$

where $< >$ designates an average over all readings. $\bar{D}$ contains information on multiple coulomb scattering but must be corrected for various errors arising from optical and mechanical imperfections of the scattering medium and measuring equipment as well as from errors of the observer.

We shall discuss the errors in the following order:—

(*a*) The temperature noise ϵ_T which is caused by differential thermal expansion of microscope parts during the measuring period.

(*b*) The reading noise ϵ_R which arises from the failure of the observer to place the goniometer hair-line accurately through the centre of gravity of the row of grains or the solid column of developed silver which defines the track in the field of view.

(*c*) The stage noise ϵ_S which arises from imperfection in the motion of the microscope stage, and focussing control.

(*d*) The grain noise ϵ_G. This can arise from the fact that, because the grains have a finite size, the centre of gravity of a finite number of grains in a row will not in general coincide with the particle trajectory. In the case of tracks consisting of solid columns of developed silver there may be other effects which cause local deviations between the centre of the silver column and the particle trajectory. One effect may be contagious development which

could lead to irregular patches, where the track is broadened. Another source of such noise is perhaps due to irregular shrinking of the emulsion in the presence of the large inhomogeneity presented by the track itself. We shall designate by grain noise any deviation between the centre of gravity of the track and the trajectory which varies in distances small compared to about $20\,\mu$.

(*a*) *Temperature Noise.*—The effect of temperature noise can best be illustrated by considering Fig. 2. The graph shows the co-ordinates

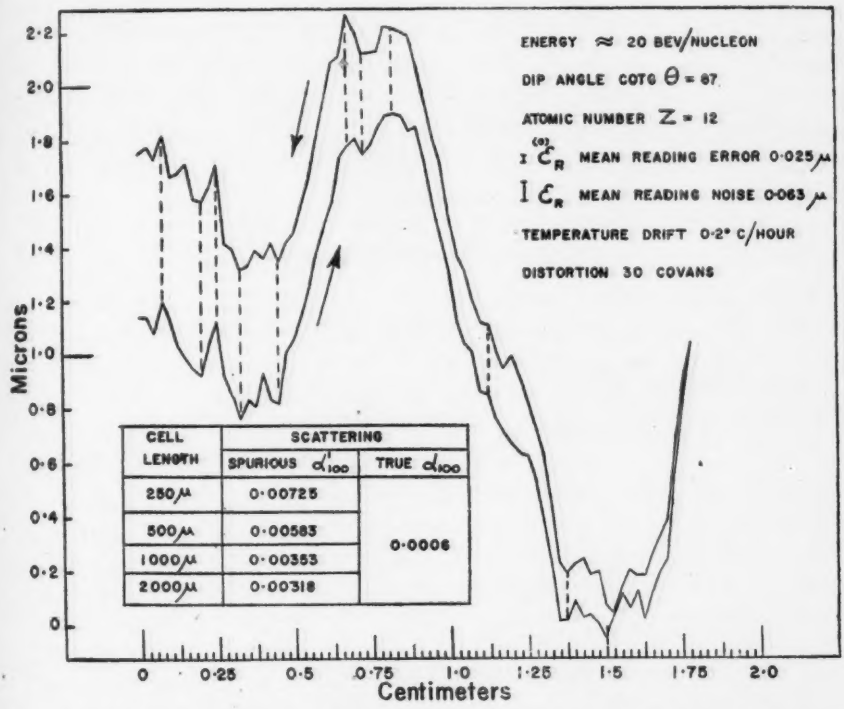

FIG. 2. The track contour due to a magnesium nucleus > 20 BeV/nucleon] measured twice in succession proceeding in the direction indicated by the arrows. The ordinate has been amplified 9,000 times.

measured at $250\,\mu$ intervals of a track in a $600\,\mu$ Ilford G-5 emulsion. The nature and energy of the particle which produced the track and other details are not relevant at this point; they are given in the figure. In order to exhibit the detailed structure of this fairly straight track the ordinate has

been enlarged 9,000 times. The lower curve has been obtained by moving along the track in the direction indicated by the arrow; the upper curve represents the results of repeating the measurements of the co-ordinates at the same points in reverse order, immediately after the first series of measurements had been completed.

The two curves shown in Fig. 2 do not coincide but show a divergence growing with the time interval which elapsed between the corresponding readings. This shift is due to a differential expansion between the arm and the stage of the Koristka microscope. The case illustrated in Fig. 2 corresponds to a differential temperature change of about 0·2° C./hour. This condition of temperature stability can be attained in an airconditioned room where all lamps are far removed from the microscope and the microscope illumination (with water filter) has been switched on several hours before taking readings. Figure 3 shows the relative motion of optical axis and stage when a 60-watt table light is switched on and later switched off at a

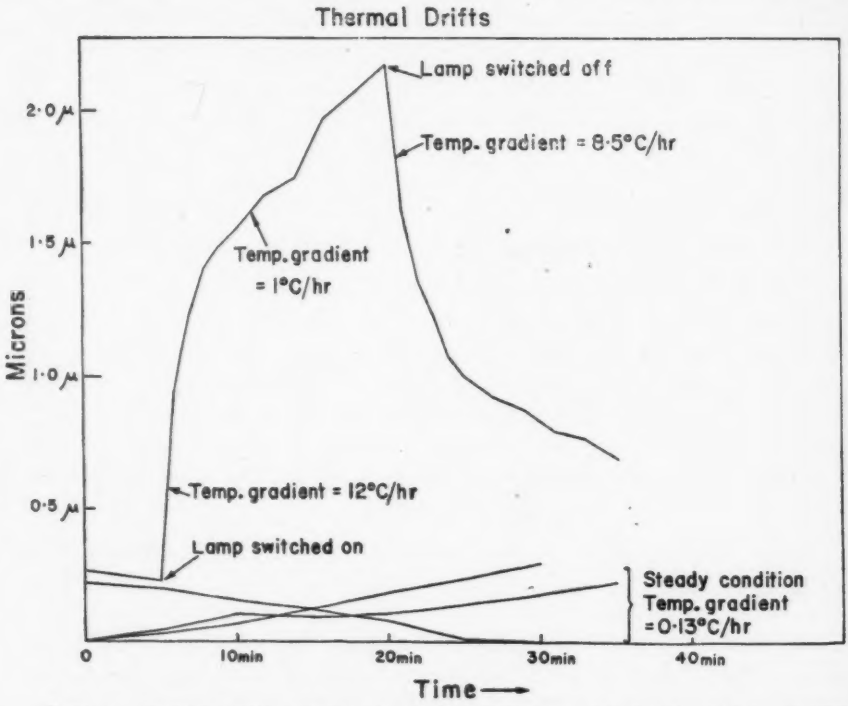

FIG. 3. Motions of the optical axis of the microscope due to changing temperature gradients.

distance of 12″ from the microscope. The differential temperature change right after switching on or off is 12° C./hour and even 15 minutes later it is still greater than ·75° C./hour.

In a Koristka microscope the temperature noise is approximately given by:

$$\epsilon_T = 3{\cdot}8 \frac{dT}{dt} \Delta t \text{ microns} \qquad (2)$$

where T is the temperature difference between microscope stage and arm in centigrade, t is the time, and Δt is the mean variation in time intervals between successive readings. Thus, even in a room which is not airconditioned, and where temperature variations between different parts of the microscope of the order of 2° C./hour occur, it is possible to make the temperature noise ϵ_T negligible by taking successive readings at regular intervals, but the regularity requirements are then fairly stringent. (Thus, for tracks of optimum thickness where as shown below, one can reduce noise from other sources to $0{\cdot}058\,\mu$, the contribution from thermal effects will dominate unless successive reading intervals differ on the average, by less than 25 seconds.)

In any case scattering measurements should not be started before the microscope has reached thermal equilibrium with its surroundings. In this investigation we made our measurements with temperature drifts of the order of 0·15° C. per hour and took readings at regular intervals differing by less than 15 seconds; under these conditions the effect of temperature drift on second differences can then be neglected.

(*b*) *Determination of Reading Noise* ϵ_R.—Returning again to Fig. 2, one sees that while the two track contours are very similar, the co-ordinates of a particular point on the track, measured at two different times, differ slightly, even when we allow for the temperature drift which has occurred during the time interval. This difference between measurements taken at the same point along the track and at the same setting of the stage, is obviously independent of the shape of track and of the stage motion; it can only arise from errors in reading the track co-ordinates. It is easy to isolate and measure this source of error even in the presence of large temperature drifts.

Let a_n and b_n designate co-ordinate values obtained for the same points on the track in the two different sets of measurements. We can subtract corresponding readings and form the second differences in order to avoid temperature effects. We then obtain for the reading *noise* ϵ_R (which is the mean absolute second difference of individual reading *errors* $^{(0)}\epsilon_R$):

$$\overline{D - D'} = \frac{1}{N}\sum_{n=1}^{N} |\,(a_n - b_n) - 2\,(a_{n+1} - b_{n+1}) + (a_{n+2} - b_{n+2})\,|$$

$$= \frac{1}{N}\sum_{n=1}^{N} |\,(a_n - 2a_{n+1} + a_{n+2}) - (b_n - 2b_{n+1} + b_{n+2})\,|$$

$$= \sqrt{2}\ \epsilon_R \qquad (3)$$

Thus, ϵ_R is the contribution of individual reading *errors* to the mean absolute value of the second difference $\bar{D}$.

More generally one can obtain from two sets of measurements of the same track the reading *noise* not only for second differences, but in all orders. In terms of the mean reading *error* ${}^{(0)}\epsilon_R$ of individual co-ordinate measurements, this noise is given by:

$$\delta_0 = <|\ a_n - b_n\ |> = \sqrt{2}\ {}^{(0)}\epsilon_R \qquad (4\,a)$$

$$\delta_1 = <|\,(a_n - b_n) - (a_{n+1} - b_{n+1})\,|> = \sqrt{4}\ {}^{(0)}\epsilon_R \qquad (4\,b)$$

$$\delta_2 = <|\,(a_n - b_n) - 2\,(a_{n+1} - b_{n+1}) + (a_{n+2} - b_{n+2})\,|>$$
$$= \sqrt{12}\ {}^{(0)}\epsilon_R = \sqrt{2}\ \epsilon_R \qquad (4\,c)$$

$$\delta_3 = <|\,(a_n - b_n) - 3\,(a_{n+1} - b_{n+1}) + 3\,(a_{n+2} - b_{n+2})$$
$$- (a_{n+3} - b_{n+3})\,|> = \sqrt{40}\ {}^{(0)}\epsilon_R \qquad (4\,d)$$

etc.*

Experimentally we should expect the zero order noise to be strongly influenced by temperature changes which occurred in the time interval between the two sets of measurements. First order differences should only be affected if the temperature changes are large and second order differences only if the temperature changes are non-linear. It is interesting to compare the theoretical noise ratios in various orders with the experimental ratios. The close agreement between the theoretical and experimental values is shown in Table I and proves not only that as expected the reading errors ${}^{(0)}\epsilon_R$ have

* The numerical factors are easily derived from the integral below: Let x_i be variables with a probability distribution $P(x_i) = \frac{1}{\sqrt{2\pi}\ x_0}\ \exp(-x_i^2/2x_0^2)$ and let a_i be constants. Then

$$<|\sum_{i=1}^{N} a_i x_i|> = \frac{1}{(2\pi x_0^2)^{N/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left|\sum_{i=1}^{N} a_i x_i\right| \exp\left(-\sum_{i=1}^{N} x_i^2/2x_0^2\right) dx_1 \cdots dx_N$$

$$= \sqrt{\frac{2}{\pi}}\, x_0 \left(\sum_{i=1}^{N} a_i^2\right)^{\frac{1}{2}}$$

a gaussian distribution but also that in our experiment temperature drifts have only a small effect on first order differences and no effect on the determination of reading noise in second or third order.

TABLE I

	δ_1/δ_2	δ_3/δ_2	δ_4/δ_2
Theoretical ..	$\sqrt{\frac{1}{3}}$ $=\cdot577$	$\sqrt{\frac{10}{3}}$ $=1\cdot825$	$\sqrt{\frac{35}{3}}$ $=3\cdot415$
Experimental ..	$\cdot659\pm\cdot04$	$1\cdot85\pm\cdot11$	$3\cdot51\pm\cdot20$

For each track we have determined the reading noise by measuring the contour twice and forming second and higher order differences of the discrepancies between the two series of measurements. The experimental values obtained in this way depend of course on the thickness of the track which is being measured. The dashed curve in Fig. 4 shows the experimentally obtained values of ϵ_R, for relativistic nuclei of various atomic numbers in Ilford G-5 emulsions. The lowest value $\epsilon_R = 0\cdot042\,\mu$ is reached for atomic numbers $Z = 4$ or 5; for lower and higher atomic number the reading noise increases. It is interesting to observe that the minimum noise corresponds to a mean reading error ${}^{(0)}\epsilon_R$ (the accuracy with which the centre of the track can be located) of only 170 Angstroms.

(*c*) *The Stage Noise* ϵ_S.—When the stage of the microscope is moved with or without simultaneous motion of the focusing fine control, the optical axis of the microscope traces a line on the emulsion surface. If this line is not straight, the motion of the microscope controls will give rise to stage noise. In order to measure this stage noise we have adopted the following procedure. The contour of a track is measured over a distance of 9 mm. using successively different sections (i) of the screw which produces the stage motion. The contour over each section is measured twice in order to determine the reading noise ϵ_R by the method described above. We chose a track due to a relativistic particle of charge $Z = 4$ for which the reading noise ϵ_R was found to be exceptionally low. We then have the following relations:—

$$\overline{D_i - D_i'} = \sqrt{2}\ \epsilon_R \qquad (5)$$

$$\overline{D_i - D_j} = \sqrt{2}\ \epsilon_{R,\,S} \qquad (6)$$

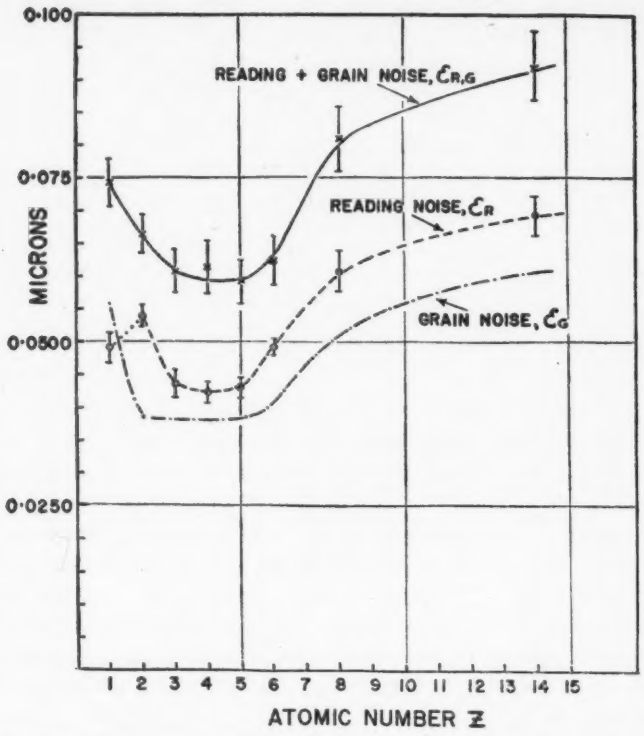

FIG. 4. The reading noise ϵ_R and the combined reading and grain noise $\epsilon_{R,G}$ in Ilford G-5 emulsions for tracks of relativistic particles with different atomic numbers.

where the first relation represents second differences obtained from two series of measurements, carried out on the same section of the screw, but at different times, while the second relation refers to second differences obtained from measurements carried out over two different sections of the screw. $\epsilon_{R,S}$ represents the combined noise due to reading error and stage motion.

We find that in our Koristka microscope ϵ_R is equal to $\epsilon_{R,S}$ within the statistical error for all values of cell lengths upto 2,000 μ and has for the particular track chosen for these measurements the numerical value 0·037 μ. We conclude that the stage noise increases our reading errors by less than 0·004 μ or 40 Å, and is negligible in all scattering measurements.

In Fig. 5, the combined reading and stage noise $\epsilon_{R,S}$ and the total noise $\epsilon_{R,G,S}$ (including grain noise) for fast particles of charge $Z = 1$ and $Z = 4$

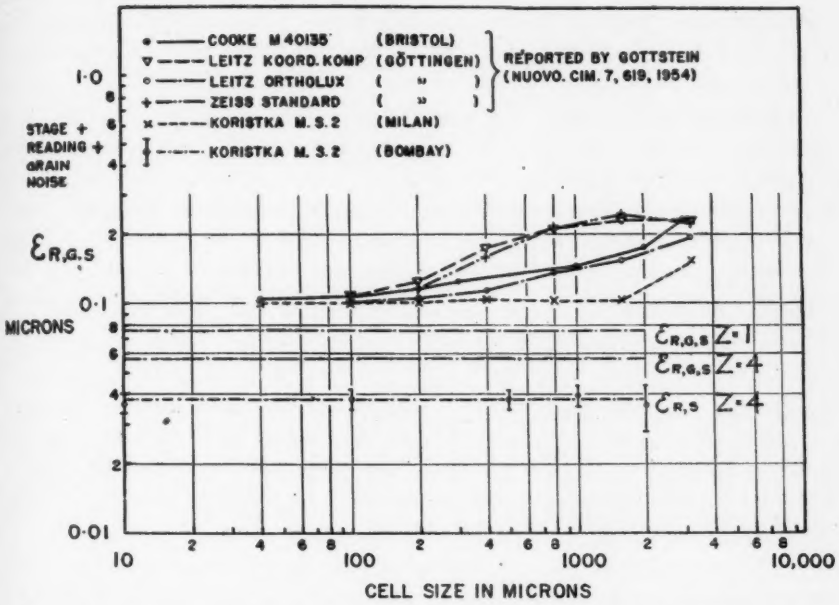

FIG. 5. The combined reading and stage noise and the total noise for tracks due to fast particles of charge Z=1 and Z=4 is measured on our Koristka microscope for various cell lengths and is compared with corresponding values of singly charged particle tracks obtained by Gottstein[14] on microscopes in other laboratories.

measured on our Koristka microscope is compared with corresponding values for singly charged particles obtained by Gottstein[14] on various other microscopes in use at Bristol, Göttingen and Milano.

(*d*) *The Grain Noise* ϵ_G.—If the tracks whose scattering is to be measured have many gaps, the position of the trajectory has to be estimated from the centre of gravity of the small number of grains which define the trajectory in the field of view. Since the centres of individual grains will be distributed in a random fashion close to the trajectory, we must expect the resulting noise to be inversely proportional to the square-root of the grain density ($\epsilon_G = A/\sqrt{n}$). Such grain noise is, therefore, much more important for scattering measurements on fast protons than for measurements on heavy nuclei. However, there are other effects which can produce localized deviations between the particle trajectory and the centre of gravity of the track, like irregular broadening by contagious development or irregular shrinkage of the emulsion in the presence of the inhomogeneity produced

by the track itself. Such effects may increase with ionization. We define, therefore, grain noise ϵ_G as the noise due to any irregularity in the distribution of developed silver around the particle trajectory provided that these irregularities occur in distances short compared to $20\,\mu$.

The following method was used for determining the grain noise. We compare again the contours of the same track obtained in two different series of measurements; however, this time the points at which the contour is determined in the second measurement series are displaced along the track by $20\,\mu$ with respect to the points used in the first series. Thereby, a different set of grains is used to determine the position of the trajectory in the two series of measurements. If we designate displacement by an asterisk we have

$$\overline{D - D^*} = \sqrt{2}\,\epsilon_{R,G} \tag{7}$$

In this way the grain noise is slightly overestimated because spurious or real scattering along the $20\,\mu$ displacements may also contribute to the difference between the two contours. But this error is negligible except for tracks whose spurious or real scattering is considerably higher than the tracks investigated here ($p\beta < 100$ MeV/c.).

The reading noise ϵ_R, the combined reading and grain noise $\epsilon_{R,G}$ and the grain noise ϵ_G (calculated from the relation $\epsilon_{R,G}^2 = \epsilon_R^2 + \epsilon_G^2$) are plotted in Fig. 4 for relativistic particles of various charges. The steep decline of ϵ_G between $Z = 1$ and $Z = 2$ is certainly an effect of the increased grain density and the resulting decrease in the mean distance between the centre of gravity of the track and the trajectory. The gradual rise of ϵ_G for higher values of Z must be attributed to one of the other causes mentioned above.

One may summarize the discussion in this section as follows:—

(1) The effect of relative thermal motions of various parts of the microscope during the period in which measurements are carried out is not in general negligible but has been rendered negligible in our experiment by the following precautions:—

(*a*) The measurements were carried out in a temperature controlled room;

(*b*) All heat sources were removed from the neighbourhood of the microscope;

(*c*) The microscope was permitted to reach thermal equilibrium after the illumination had been switched on; and

(*d*) Successive readings were taken at regular intervals.

(2) Stage noise in our Koristka microscope is too small to affect measurements even for those tracks which permit the optimum accuracy.

Under these conditions we can write:

$$\bar{D} = (\epsilon_{R,G}{}^2 + \Delta^2)^{\frac{1}{2}} \tag{8}$$

where as before:

$\bar{D}$ is the mean absolute value of the second co-ordinate differences;

$\epsilon_{R,G}$ is the combined reading and grain noise; and

Δ contains the contribution of spurious and real scattering to the measured second differences.

For singly and doubly charged particles, and for particles with atomic number $Z \geqslant 8$ where the grain noise is large, we always repeat the contour measurements using a displacement of $20\,\mu$ as explained before. One obtains then from Equations 7 and 8:

$$\Delta^2 = (\bar{D})^2 - \tfrac{1}{2}(\overline{D-D^*})^2 \tag{9}$$

For tracks due to particles of charge $2 < Z < 8$ we sometimes used the same procedure but more often we repeated the contour measurements without displacement and used for ϵ_G the constant value $0{\cdot}038\,\mu$ which as one can see from Fig. 4, does not introduce measurable errors. This latter procedure leads to:

$$\Delta^2 = (\bar{D})^2 - \tfrac{1}{2}(\overline{D-D'})^2 - (0{\cdot}038)^2 \tag{10}$$

Equations 7, 9 and 10 are equally valid if one uses third or higher order differences rather than second differences, provided the constant 0·038 is multiplied by the appropriate factors ($\sqrt{20/6}$, $\sqrt{70/6}$, etc.). The use of third differences eliminates from scattering measurements most of the effects of ordinary large-scale emulsion distortions.[15]

III. Experimental Procedure and Results

The following types of tracks were selected for measurement:

Multiply (and a few singly) charged particles whose energy from subsequent nuclear interactions was known to exceed 5 BeV/nucleon;

Heavy primary nuclei incident at geomagnetic latitudes $\lambda \leqslant 19°$.

Since it has been shown previously[2] that all heavy primary particles arriving at a given latitude have energies in excess of the geomagnetic cut-off energy (the energy required for bare nuclei approaching the earth's orbit) both criteria exclude particles below 5 BeV/nucleon.

Sixty tracks have been obtained from a block of 600 μ stripped emulsions flown from Delhi in 1954. The remaining 40 tracks came from many different types of nuclear emulsions. These include two additional blocks of stripped emulsions flown and processed at different times, and glass backed emulsions with emulsion thicknesses varying from 250 μ to 1,000 μ. Among the latter were plates kindly supplied to us some time ago by Dr. Camerini, and others by Dr. Anand, which had been processed at the H. H. Wills Physical Laboratory at Bristol, as well as plates processed at the University of Chicago and at the University of Melbourne, for which we want to express our thanks to Professor M. Schein and Dr. V. D. Hopper, respectively.

The geometric selection criteria were based, not on the length of the tracks in the emulsion, but on their angle of dip. Comparable numbers of tracks were selected with dip angle values in the four intervals:

$$8 < \text{cotg}\,\theta \leqslant 12$$

$$12 < \text{cotg}\,\theta \leqslant 20$$

$$20 < \text{cotg}\,\theta \leqslant 30$$

$$30 < \text{cotg}\,\theta$$

The contour of each track was measured twice in succession using 250 μ steps. If the tracks were due to singly or doubly charged particles or to a particle of charge $Z \geqslant 8$, the two series of measurements were shifted by 20 μ with respect to each other and the scattering was determined from Equation 9. For tracks of other particles both series of measurements were in most cases carried out with the same setting and the scattering determined from Equation 10. The scattering value was calculated using second and third order co-ordinate differences. In some cases third order differences gave lower scattering than second order differences, probably because of S or C-shaped distortion. In those cases the value was calculated from third differences. These cases were mostly confined to the groups with dip angle $\text{cotg}\,\theta < 20$.

Since all tracks investigated here are very long, statistical fluctuations are of no great importance in the results presented below. The consistency of measurements on individual tracks can best be judged from results obtained on the same track by repetition of measurement. An example chosen at random is given in Table II.

TABLE II

Results of Independent Scattering Measurements on a Track due to a Particle of Charge $Z = 12$

Cell length	$\bar{D}$-values (μ)			
	250 μ	500 μ	750 μ	1,000 μ
1	0·0805	0·1300	0·1875	0·2325
2	0·0845	0·1265	0·1910	0·2320

In Fig. 6 we have averaged separately for each group of dip angles the mean scattering value Δ and plotted it against cell length. The same

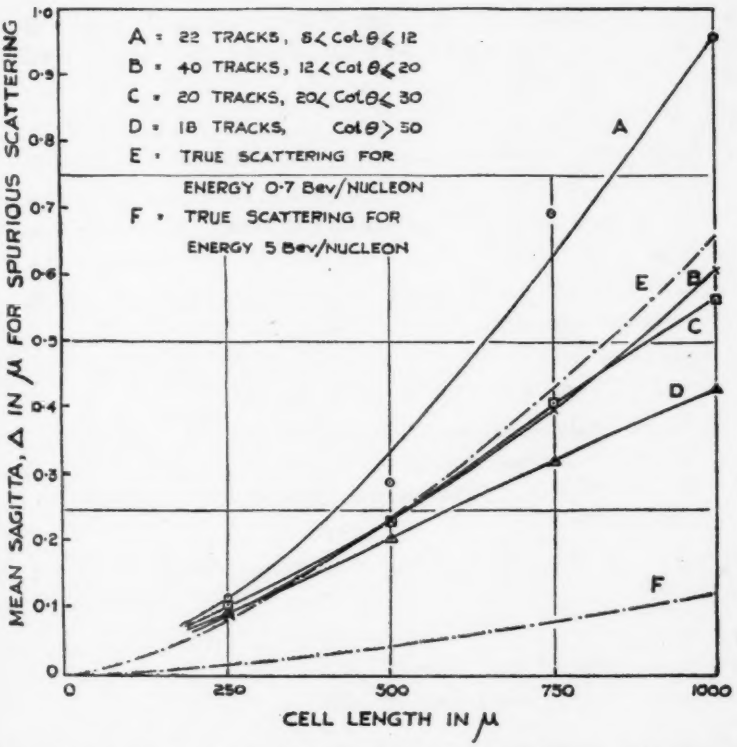

FIG. 6. The second order co-ordinate differences due to spurious scattering are plotted as a function of cell lengths. Results have been averaged for all tracks belonging to the indicated interval of dip angles. The dotted curves show the co-ordinate differences expected from true multiple coulomb scattering of heavy nuclei with energies 0·7 and 5 BeV/nucleon respectively.

graph shows the expected true scattering value Δ_t for heavy nuclei of energy 0·7 and 5 BeV/nucleon.

From Fig. 4 and Fig. 6 one can see that the average spurious scattering greatly exceeds noise values and that spurious scattering increases with cell length as rapidly as true scattering. It becomes also apparent that on the average flatter tracks have somewhat lower spurious scattering than steeper tracks in the interval of dip angle 1·9–7°, but below 1·9° we find that spurious scattering becomes fairly independent of dip angle.

Although Fig. 6 exhibits the general trend of spurious scattering with track inclination and with cell length, it does not yet convey an idea of the spread of values on individual tracks. This is shown in Fig. 7 *a* and *b*. The abscissa in these figures is the experimental scattering value α'_{100} as marked on top of the graph and the corresponding energy per nucleon for heavy primaries marked on the bottom.† The ordinate shows the distribution in dip angles. (All tracks with cotg $\theta > 65$ have been plotted at cotg $\theta = 65$.)

The graphs are divided into three regions marked Z, Z − 1 and Z − 2 respectively. These regions have the following significance:

If we had no *a priori* knowledge that the track under investigation was caused by a relativistic particle of energy above 5 BeV/nucleon; and

if we, therefore, attributed the measured scattering to true multiple coulomb scattering; and

if we tried to determine the atomic number of the corresponding nucleus from δ-ray density and scattering;

Then we would obtain the correct charge value for those tracks which fall into the region Z; the charge of those which fall into regions Z − 1 and Z − 2 would be underestimated by one and two units, respectively.

Figure 7 *a* shows the distribution of scattering results using 250 μ and 500 μ cells; Fig. 7 *b* shows the results obtained with 750 μ and 1,000 μ cells.

On the abscissa we have indicated the limits set to the measurements by the various forms of noise discussed in Section II, as well as the energy limit imposed by the geomagnetic cut-off. In the absence of spurious

† We have used the scattering constant K given by Gottstein *et al.* [16] Since we are dealing with spurious, not real, scattering we have not used any large angle cut-off, but the effect of large angle deflection is in any case unimportant.

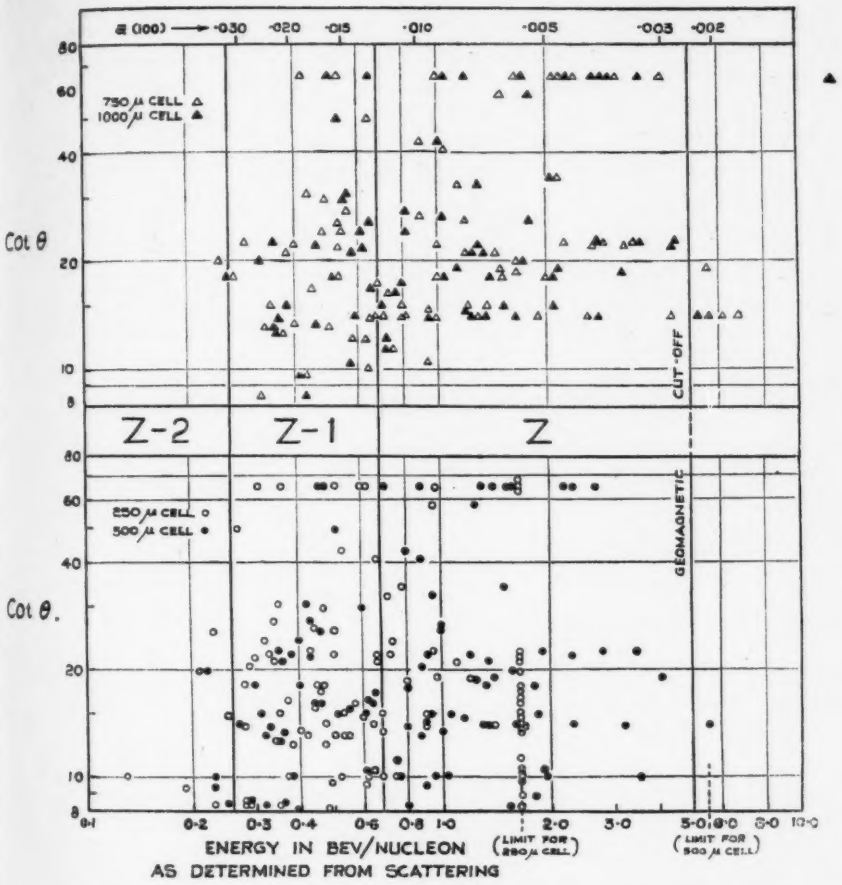

FIG. 7. Each point gives the dip angle, the spurious scattering parameter α'_{100} and the "energy" derived from it, for one of the tracks investigated. Different symbols refer to different cell lengths employed. The abscissa also shows the highest energy measurable with 250μ and 500 μ cells in the presence of noise. The regions marked Z, Z − 1, Z − 2 contain the tracks for which the atomic numbers derived from ionization *plus* scattering measurements is underestimated by zero, one or two units respectively. The true multiple coulomb scattering for all 100 tracks should place each point to the *right* of the line marked geomagnetic cut-off.

scattering the points in Fig. 7 *a* should all lie on the line determined by the energy limit set by noise and in Fig. 7 *b*, all points should lie to the right of the energy limit imposed by the geomagnetic cut-off. Almost all points, however, lie far to the left of these limits.

Together Fig. 7*a* and *b* show that spurious scattering will dominate completely over other sources of noise for tracks due to particles of a few hundred MeV/nucleon. The effect is somewhat more pronounced for steeper tracks and shorter cell lengths but, for dip angles below 1·9° (cotg $\theta = 30$) and cell lengths above 500 μ, it becomes nearly independent of these variables.

The graph furthermore indicates that scattering values corresponding to energies of 7 BeV/nucleon for heavy nuclei or 14 BeV for protons are occasionally obtainable with cell lengths as low as 1,000 μ.

In Table III we have listed details of measurements on those tracks which produce nuclear collisions indicating energies in excess of 50 BeV/nucleon, that is ten times higher than the minimum value accepted in this

TABLE III

Spurious Scattering of Tracks of Particles with Energies Above 50 BeV/Nucleon

Description of the plate	Atomic No. Z	True Energy E_T BeV/nucleon	Reading+Grain Noise ϵ_{RG} (μ)	Spurious Scattering $\Delta(\mu)$ for cell lengths			Expected True Scattering Δ_t
				250 μ	500 μ	1000 μ	1000 μ
N-stack (600 μ stripped)	2	2000	·058	·070	·121	·264	·00037
Melbourne (400 μ g.b.)	1	1000	·088	< ·088	·169	·099	·00149
N-stack (600 μ stripped)	2	300	·058	·118	·416	1·370	·0025
			·072	< ·040	·100	·220	·0025
			·074	·058	·073	·135	·0025
Melbourne (400 μ g.b.)	1	150	·081	·071	·170	·338	·0100
K-stack (600 μ stripped)	6	90	·077	·107	·324	·775	·0083
			·059	·045	·086	·213	·0083
			·063	< ·040	·137	·386	·0083
			·068	< ·040	< ·040	·114	·0083
			·061	·062	·133	·329	·0083
			·051	< ·040	·070	·408	·0083
N-stack (600 μ stripped)	6	60	·055	·055	·141	·304	·0127
			·073	·139	·360	·815	·0127
N-stack (600 μ stripped)	5	50	·061	·118	·446	1·221	·0150
Bristol (400 μ g.b.)	6	50	·071	·049	·052	·260	·0150

experiment. It appears that results here are not significantly different from the rest of the measurements given in Fig. 7 *a*, *b*. But the data in Table III show clearly that because spurious scattering in individual tracks increases with cell length (t) almost like $t^{3/2}$, it cannot be distinguished from true multiple coulomb scattering by comparing results for various cell lengths.

Since the increase of spurious scattering with cell length does not permit us to distinguish true from spurious scattering, we need additional information about a particle before we can determine, or obtain limits on, the product of its velocity and momentum from scattering measurements. There are three commonly occurring cases of additional information.

It may be known from range or variation of ionization density along the track that a particle was slow and that the expected true multiple coulomb scattering should be much higher than scattering values measured on known high energy tracks; we can then neglect spurious scattering and make the approximation $\Delta_{obs.} \approx \Delta_{true}$. This approximation applies in many experiments which have been carried out with nuclear emulsions in the past, in particular in the experiments designed to determine the mass of slow K-mesons.

It may be found that there is a second track parallel and close to the track under investigation. We shall only discuss the simplest and most common case, namely when the two tracks are due to particles of equal mass and energy. This applies to tracks found in cyclotron exposed emulsions and it applies to the showers of relativistic α-particles which represent the evaporation of a partially destroyed heavy primary nucleus of great energy. Such a case was illustrated in Fig. 1. We have then the following relations:

$$\overline{D_1 - D_2} = \sqrt{2}\,(\epsilon_{R,G}{}^2 + \Delta_t{}^2)^{\frac{1}{2}} \tag{11}$$

$$\left(\overline{D_1 - D_1{}^*}\right) = \sqrt{2}\ \epsilon_{R,G} \tag{7}$$

therefore

$$\Delta_t{}^2 = \tfrac{1}{2}\left[\left(\overline{D_1 - D_2}\right)^2 - (D_1 - D_1{}^*)^2\right] \tag{12}$$

D^* as before refers to co-ordinate readings displaced by $20\,\mu$ with respect to those which determine the second difference D. The subscript 1 and 2 refer to the two tracks, and Δ_t contains only true and no spurious scattering.

The well-known result, namely that relative scattering on closely spaced tracks is more reliable than measurements on individual tracks, is examplified by the tracks whose contour was given in Fig. 1. The measured values are listed in Table IV.

TABLE IV

Individual and Relative Scattering Measurements on Two Parallel α-Particle Tracks

(Expressed in microns)

Cell length	250 μ	500 μ	1000 μ	2000 μ
$\overline{D_1}$	·248	·540	·937	1·31
$\overline{D_1^*}$	·247	·527	·942	1·33
$\frac{\overline{D_1 - D_1^*}}{\sqrt{2}} = \epsilon_{BG}$	·077	..	..	..
Δ_1	·237	·53	·94	1·31
$\overline{D_2}$	·247	·545	·994	1·33
$\overline{D_2^*}$	·258	·562	1·02	1·32
$\frac{\overline{D_2 - D_2^*}}{\sqrt{2}} = \epsilon_{BG}$	·066	..	..	..
Δ_2	·239	·54	·99	1·32
Δ_t	·048	·080	·0715	·072

Table IV shows that the scattering measurements are highly reproduceable even after shifting the track by 20 μ. It also shows that the two tracks give identical results, proving again that reading errors and grain noise are unimportant.

The relative scattering Δ_t of the two tracks does not exceed the noise value for cell lengths as high as 2,000 μ. The corresponding energy of the α-particles taking account of statistical uncertainties is $>$ 20 BeV/nucleon in fair agreement with the energy of 15 BeV/nucleon estimated from the angular spread of the α-particle fragments emanating from this collision. On the other hand the scattering value, obtained on each track separately (which includes spurious scattering), corresponds to very low energies. It should be emphasized again that this event occurs in an emulsion which throughout has fairly moderate distortion ($<$ 50 covans).

The third case of additional *a priori* information applies to all tracks studied in our experiment; we know that because of our selection criteria, true scattering must be much smaller than the observed scattering values. Therefore:

$$\Delta_{obs.} \approx \Delta_{spurious}$$

We can, therefore, get some information on the nature of the displacements responsible for spurious scattering by considering the statistical distribution of observed D-values. We shall not in this paper make a detailed investigation of the type of displacement which is responsible for the effect of spurious scattering. Such an investigation is most conveniently carried out with plates which are exposed in the Bevatron and will contain many parallel tracks of mono-energetic high energy particles. Some tentative information on the nature of the dislocations has, however, been deduced from our data in the Appendix.

IV. Spurious Scattering in its Relation to the Charge and Energy Spectrum of Primary Cosmic Rays

We have shown that the determination of multiple coulomb scattering in emulsions on tracks of high energy particles is strongly affected by small-scale dislocations which occur during processing and whose effect on tracks is indistinguishable from true scattering.

At present the range of validity of scattering measurements is limited by this spurious scattering and not, as has been generally believed, by various other types of noise and errors of measurement. Statistically as can be seen from Fig. 7, the spurious scattering parameters α'_{100}, determined for many tracks, have a broad distribution ranging from low values up to a maximum at a fairly sharply defined upper limit. This limit lies between ·025 and ·030°; it decreases slightly with increasing cell length and increases with increasing dip angles above $\sim 2°$. The occasional occurrence of very low scattering values does not contradict these findings. Spurious scattering values appreciably larger than the statistically determined upper limit do also occur occasionally.

Spurious scattering is in our opinion responsible for the marked difference between the *energy spectra of primary nuclei* obtained by Dainton, Fowler and Kent,[9] by Gottstein,[10] and by Waddington[12] all of whom used *ordinary* scattering measurements and the spectra obtained by Kaplon, Peters, Reynolds and Ritson[2] who used *relative* scattering measurements, nuclear interactions, the latitude effect and the range of heavy nuclei in absorbers. The measurements which we have discussed in the previous section show that only a small fraction of the relativistic primaries can be identified as relativistic by ordinary scattering measurements.

For instance, there is general agreement about the total number of primary nuclei with atomic number $Z \geqslant 6$ incident at high latitude (geomagnetic latitude 54–55°). But there is no agreement on what fraction of these nuclei carry energy in excess of 1·6 BeV/nucleon. Experiments not

involving scattering measurements are consistent with each other; they include investigations with nuclear emulsions[2, 17] rocket-borne ionization chambers,[5] cloud chambers triggered by a proportional counter telescope,[7] and counter telescopes which include a Czerenkov detector.[8] These experiments yield a flux of heavy particles with energy $\epsilon > 1 \cdot 6$ BeV/nucleon which is about three times as large as the flux determined from scattering measurements in emulsion.[9, 10, 12]

The charge determination on heavy primary nuclei can also not be carried out safely in experiments which involve multiple scattering measurements in emulsion. Table V shows the percentage of high energy particle tracks for which spurious scattering is so large ($\alpha'_{100} > \cdot 0125°$) as to lead to an underestimate of charge by one or more units. The percentage of wrong charge assignments has been given separately for tracks of various inclination and for various cell lengths.

TABLE V

Dip angle cotg θ / Cell lengths	Percentage of underestimated charges derived from scattering measurements			
	250 μ	500 μ	750 μ	1000 μ
8 to 12	64%	43%	60%	75%
12 to 20	50%	47%	39%	30%
20 to 30	60%	42%	48%	40%
>30	47%	22%	25%	25%

In the experiment of Dainton, Fowler and Kent,[9] who attempted to determine the primary charge distribution from measurements of scattering and of δ-ray density, nearly 80% of the tracks fall into dip angle interval $12 < \text{cotg}\,\theta < 20$ and were scattered with 250 μ, to 500 μ cells. Other investigators have used still steeper tracks and shorter cell lengths for similar measurements. Figure 7 *a* and Table V show that under these conditions reliable charge determinations cannot be obtained, and that the charge will be underestimated in 50–60% of the cases.

In the presence of a strong Carbon peak in the primary radiation this must lead to a large spurious peak of Boron nuclei, and thereby to an

over-estimate of the relative intensity of the light elements Li, Be, B compared to the C, N, O group.

Apart from the presence of spurious scattering in the measurements, there are other grounds on which the conclusion of these authors, namely that the primary intensity of Li, Be, and B is comparable to that of C, N and O, can be criticised. The number of Lithium nuclei was corrected for an assumed lack of detection efficiency and was raised from the observed value of 45 to the corrected value of 130. The detection efficiency had not, however, been measured directly but was inferred from the observed Boron flux, which in turn, as we have seen, must to a large extent be spurious.

Another reason for the large primary Li, Be, B flux obtained by these authors, lies in the method adopted in correcting for the Li, Be, B nuclei produced by fragmentation of heavier primaries in the upper atmosphere. The fragmentation probability was taken from measurements carried out with relativistic particles,[1] but it was applied to the mostly non-relativistic nuclei incident at high latitude. Since it is known that the mean free path for nucleons in nuclear matter varies by factors of order 2 or 3 from energies of 0·3 BeV/nucleon to very high energies (the maximum occurring near ·6 BeV),[18] it seems reasonable to expect different fragmentation probabilities at different latitudes. By studying the fragmentation of heavy primaries in substances of low atomic weight it has been shown[19] that the probability for obtaining Li, Be, B fragments is fairly high at medium and high latitudes and that the number of Li, Be, B nuclei produced by fragmentation at balloon altitudes is much larger than was assumed by Dainton *et al*.

Since the charge distribution cannot be reliably determined by scattering measurements, there remains at present only one method which has been suggested for studying the charge distribution in the *low energy part* of the spectrum. This method consists in accurate measurements of ionization *versus* range and is confined to tracks of particles which actually come to rest in nuclear emulsion; it has been applied by Waldeskog,[20] and has so far given no evidence for a large Li, Be, B flux. But more work is needed to improve statistics in this experiment.

Measurements in the *high energy part* of the spectrum are easier to make and methods are more numerous.

Bradt and Peters,[1] who used a stack of glass backed emulsions with alternating high and low sensitivity, obtained at geomagnetic latitude $\lambda = 30°$ as an upper limit for the ratio R_0 between the primary flux of Lithium, Beryllium, Boron nuclei and that of Carbon, Nitrogen and Oxygen nuclei

at the top of the atmosphere the value $R_0 \leqslant 0{\cdot}1$. Their experiment has recently been repeated and extended to include primaries of charge $Z = 2$.[4] This last investigation includes a check of detection efficiency, which was found to exceed 85% for tracks with grain density corresponding to four times minimum ionization (primary α-particles) and must, therefore, be very close to 100% for tracks due to particles of charge $Z \geqslant 3$.

Stix[7] in two very beautiful and difficult experiments identified particles at balloon height and geomagnetic latitude $\lambda = 41°$ by the pulse height in three proportional counters and at the same time by tracks in a cloud chamber triggered by these counters. He obtains $R_0 = 0{\cdot}05\,^{+0{\cdot}32}_{-0{\cdot}05}$. The upper limit is calculated by including dubious cases as belonging to the Li, Be, B group and by using not the standard deviation but the 95% probability for the result to fall within the limits.

Webber[8] at the same latitude, used a Czerenkov counter within the solid angle of a Geiger counter telescope. He obtains at *balloon altitude* a ratio Li Be B/$Z \geqslant 6 = 0{\cdot}16$. Correction for light nuclei produced by fragmentation above the balloon will reduce this ratio, but the statistical errors are large and uncertain.

Existing evidence at latitudes up to 41° favours, therefore, a low intensity of Li, Be and B nuclei. Since it is known that interstellar gas contains mostly hydrogen and that collisions of heavy nuclei ($Z \geqslant 6$) with hydrogen lead to disintegrations in which frequently the incident nucleus suffers only partial destruction, some Li, Be, B must be created in such collisions and must, therefore, be present in the primary beam. The observed intensity sets an upper limit to the fraction of heavy primaries which have suffered collisions in interstellar space and, therefore, to the amount of matter which they have traversed since their acceleration.

In order to calculate this amount of matter one has to consider separately the fragmentation of the C, N, O group, and that of heavier primaries, because the processes by which Lithium, Beryllium and Boron are produced in collision with hydrogen are basically different, in the two cases.[21] The evaporation of a nuclear fragment of charge $Z \geqslant 3$ in a collision between a heavy nucleus and hydrogen is rare, but its complete disintegration into singly and doubly charged particles is also very rare and the most common collision results in a reduction of charge of the heavy nucleus. On the other hand the collision of C, N, O nuclei with hydrogen must lead to the production of Li, Be and B residues except in those cases where fragmentation into singly and doubly charged fragments is complete.

The probability (p) of producing a Li, Be, B fragment in a collision between a nucleon and a carbon, nitrogen or oxygen nucleus has been measured by various investigators.

Harding[22] by applying charge conservation to the prongs of cosmic ray stars in gelatine found that in 70% of the collisions the residue must have atomic number $Z > 2$. (Hodgson[23] who carried out an experiment similar to Harding's, obtained a value for p which is certainly much too low because it cannot be corrected for the large number of one and two pronged stars missed in his experiment.)

The method of measuring the probability p used by Noon, Kaplon and Ritson[24] seems preferable for our purpose because it is based on collisions of the same type as those presumed to occur in interstellar space. Out of 87 nuclear interactions produced in an emulsion block detector by the heavy primary component incident at geomagnetic latitude $\lambda = 41°$, they selected those due to collisions with protons characterised by the absence of evaporation tracks, by a recoil proton at much larger angle than the fast fragments of the incident nucleus, and by the fact that charge balance can be obtained on the assumption that the collision occurred with hydrogen in the target material. They obtained:

$$p = 0{\cdot}8 \pm 0{\cdot}5$$

It would be very desirable to reduce the statistical error by further observations. Their result, however, is in rough agreement with that of Harding.

If we assume $p = 0{\cdot}7$ and remember that the group comprising C, N, O nuclei dominates the production by collision of Li, Be, B, one obtains an upper limit for the amount of interstellar hydrogen gas traversed by the primary cosmic radiation equal to 14% of the interstellar mean free path of C, N, O nuclei. We may take 300 millibarns as the cross-section for the interaction of fast proton with nitrogen and obtain then an upper limit of $0{\cdot}75$ g./cm.^{2} for the amount of interstellar material which the primary cosmic radiation has traversed in the time interval which elapsed between acceleration and arrival at the outer atmosphere. In view of the statistical uncertainty in the parameter p a more refined calculation does not as yet seem justified. This upper limit corresponds to the amount of matter encountered in approximately a thousand traversals of the thickness of the galactic disc.

One can obtain a different upper limit by the following considerations. Iron is the heaviest element definitely identified among the primary particles. Iron nuclei constitute $\frac{1}{2}$% of the primary radiation in a given velocity interval

(at a given geomagnetic latitude the percentage is larger) and carry 2% of the nucleons. If one makes the extreme assumption that at the time of acceleration the primary beam contained only iron nuclei one obtains from the fact that 2% did not suffer destructive collisions with interstellar hydrogen an upper limit of four mean free paths or 9·1 g./cm.^{2} of hydrogen. If on the other hand one makes the more plausible assumption that at the time of acceleration the ratio of Iron to Carbon nuclei was the same as the abundance ratio of these elements in the universe (approximately 0·2) one obtains an upper limit of 0·7 mean free paths or 1·6 g./cm.^{2} of hydrogen, not much higher than that derived from the abundance minimum in the primary Lithium, Beryllium and Boron component.

Thus, it seems safe to conclude from the chemical composition of primary nuclei, that few if any of the primary particles approaching our atmosphere have suffered collisions in the interstellar space.

APPENDIX I

Figure 1 shows clearly that the *magnitude* of the dislocations in processed emulsions often varies appreciably in distances as small as 250 μ. If the *direction* of the displacements also changes within such small distance, one might expect to obtain a gaussian distribution G (D) for the individual values of the second differences D. On the other hand if the direction of displacements persist over larger distances comparable with the length of the tracks themselves, one must expect a quite different distribution. In that case, the resulting D-values will be proportional to the sine of the angle ψ between the track and the displacement vector and if this angle is arbitrary, the distribution of D-values takes the form:

$$F(D) = \frac{1}{\sqrt{2\pi}\, x_0} \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{\exp\left[-D^2/2x_0^2\{1+(x_1^2/x_0^2)\sin^2\psi\}\right]}{\sqrt{1+(x_1^2/x_0^2)\sin^2\psi}}\, d\psi \quad (14)$$

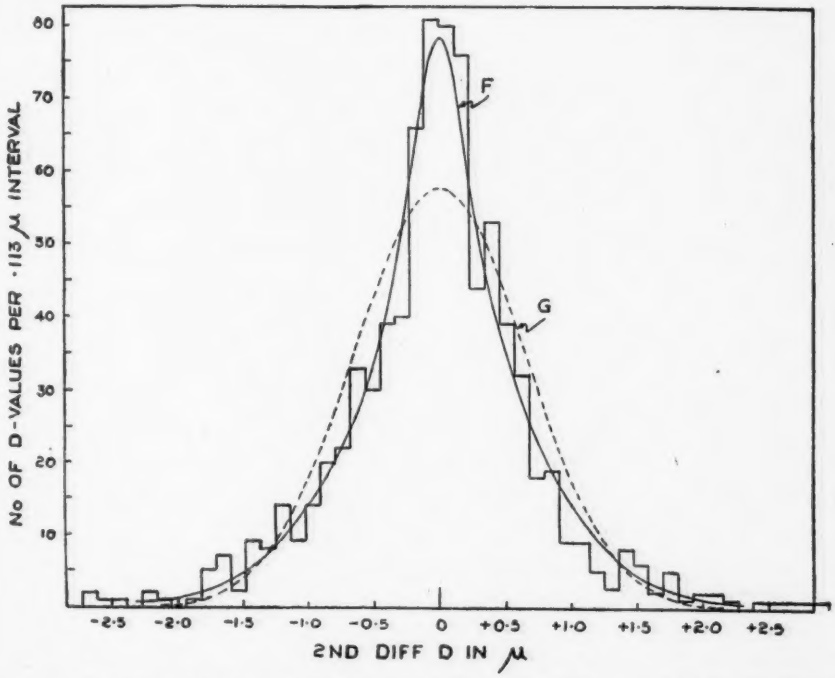

FIG. 8. The distribution of D-values (second order differences due to spurious scattering) is compared with two theoretical curves (see text).

Here x_0 represents the contribution of ordinary noise to the spread of D-values and has been taken as $x_0 = 0{\cdot}075$. In Fig. 8, we have plotted a histogram of 826 D-values obtained with 1,000 μ cells on tracks whose dip angle is less than $1{\cdot}9°$. The gaussian distribution and the distribution F (D) have also been plotted. The distribution curves are normalized to the same area as the histogram and the parameter $x_1/x_0 = 12{\cdot}3$ has been chosen so as to yield the experimental value for the mean of the absolute second difference, namely $\bar{D} = 0{\cdot}513\,\mu$.

Since F (D) fits the experimental distribution while G (D) does not fit, we may tentatively conclude, that the dislocations which give rise to spurious scattering change in magnitude over short distances but persist in direction for much longer distances. If this conclusion is correct, we should find in the same emulsion tracks with low and tracks with high spurious scattering.

This seems to be borne out by our data. In Fig. 9, we have plotted the energies per nucleon obtained with 1000 μ cells for some tracks which

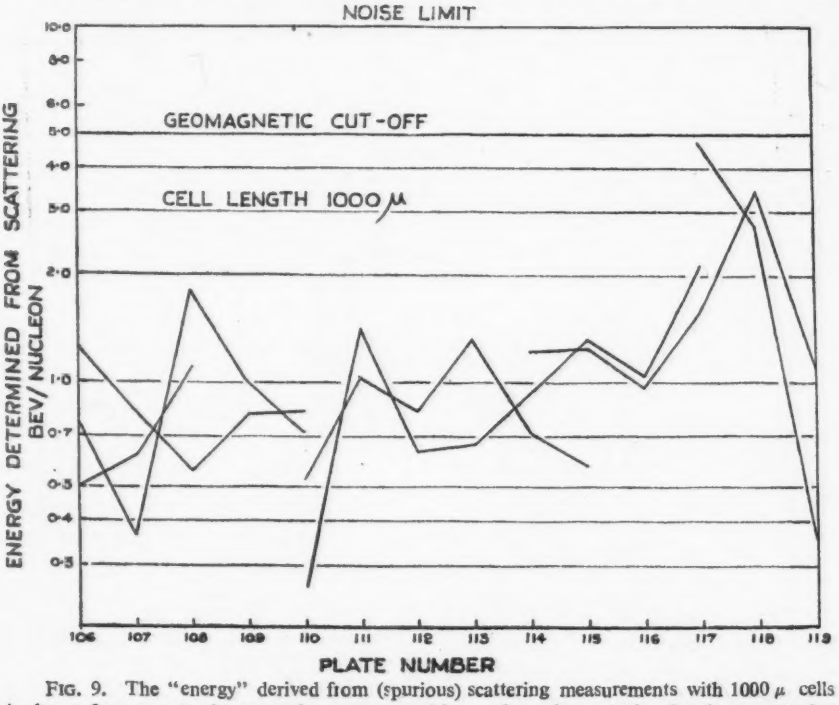

FIG. 9. The "energy" derived from (spurious) scattering measurements with 1000 μ cells is shown for some tracks traversing many emulsions of a given stack. Spurious scattering can differ widely for two different tracks traversing the same emulsions sheet and for the same track traversing different emulsion sheets.

traverse several emulsion sheets. The graph indicates that energy values for the same particle are quite erratic from plate to plate, and that a single emulsion contains in general tracks which have large spurious scattering as well as tracks which have comparatively low spurious scattering. This we interpret tentatively as connected with the angle between the track and the dislocation vectors.

Finally, if the spurious scattering is proportional to the sine of the angle ψ between track and dislocation vector, and if these angles are arbitrarily distributed, one expects for a large number of tracks a certain probability distribution for values of α'_{100} namely:

$$P(\alpha)\, d\alpha = \frac{2}{\pi} \frac{d\alpha}{\sqrt{\alpha_0^2 - \alpha^2}}$$

where α_0 is determined by the observed upper limit of α'_{100}. We have only 17 tracks with dip angle cotg $\theta > 30$ which have been scattered with 1000 μ cell length (Fig. 7 *b*) and can, therefore, make only a rough comparison between expected and observed distribution of spurious scattering values. We find:

	Predicted	Observed
$\alpha'_{100} > \cdot0150$ ..	0	1
$\cdot0150 > \alpha'_{100} > \cdot0110$ ..	7·8	4
$\cdot0110 > \alpha'_{100} > \cdot0075$ ..	3·6	4
$\cdot0075 > \alpha'_{100} > \cdot0037$ ..	2·9	5
$\cdot0037 > \alpha'_{100}$ ..	2·7	3

References

1. H. L. Bradt and B. Peters .. *Phys. Rev.*, 1950, **80**, 943.
2. M. F. Kaplon, B. Peters, H. L. Reynolds, and D. M. Ritson *Ibid.*, 1952, **85**, 295.
3. Y. Eisenberg .. *Ibid.*, 1954, **96**, 1378.
4. B. Peters .. *Proc. Ind. Acad. Sci.*, 1954, **40**, 230.
5. J. A. Van Allen .. *Phys. Rev.*, 1951, **84**, 791.
6. G. J. Perlow, L. R. Davis, C. W. Kissinger, and J. D. Shipman (Jr.) *Ibid.*, 1952, **88**, 321.
7. T. H. Stix .. *Ibid.*, 1954, **95**, 782.

8. W. R. Webber. .. *M.Sc. Thesis*, 1955, Iowa State University.
9. A. D. Dainton, P. H. Fowler, and D. W. Kent *Phil. Mag.*, 1952, **43**, 729.
10. K. Gottstein .. *Ibid.*, 1954, **45**, 347.
11. M. F. Kaplon, J. H. Noon and G. W. Racette *Phys. Rev.*, 1954, **96**, 1408.
12. C. J. Waddington .. *Phil. Mag.*, 1954, **45**, 1312.
13. P. H. Fowler .. *Ibid.*, 1950, **41**, 169.
14. K. Gottstein .. *Nuovo Cim.*, 1954, **7**, 619.
15. S. Biswas, E. C. George, B. Peters and M. S. Swamy *Ibid.*, Suppl. Ser. 9, 1954, **12** (2), 369.
16. K. Gottstein, M. G. K. Menon, J. H. Mulvey, C. O'Ceallaigh and O. Rochat *Phil. Mag.*, 1951, **42**, 708.
17. M. F. Kaplon, G. W. Racette and D. M. Ritson *Phys. Rev.*, 1954, **93**, 914.
18. R. W. Williams .. *Ibid.* (In press)
19. J. H. Noon and M. F. Kaplon .. *Ibid.*, 1955, **97**, 769.
20. B. Waldeskog .. *Arkiv. for Fysik*, 1954, **7**, 475.
21. L. Marquez and I. Perlman .. *Phys. Rev.*, 1951, **81**, 953.
22. J. B. Harding .. *Nature*, 1949, **163**, 440,
23. P. E. Hodgson .. *Phil. Mag.*, 1951, **42**, 955.
24. J. H. Noon, M. F. Kaplon and D. M. Ritson *Phys. Rev.*, 1953, **92**, 1585.

468-55 Printed at the Bangalore Press, Bangalore City, by C. Vasudeva Rao, Superintendent, and Published by The Indian Academy of Sciences, Bangalore.

NOTICE TO AUTHORS

Scientific papers intended for publication in the *Proceedings of the Indian Academy of Sciences* can be accepted only when they are communicated by a Fellow of the Academy whose duty shall be to satisfy himself that such communications are fit to be read at the Meeting of the Academy and published in its *Proceedings*.

Papers should not ordinarily exceed fifty pages of foolscap. MSS. should be either typewritten or written in legible hand on one side of the paper. All papers should be carefully revised by the authors and should be absolutely in final form for printing. Position for text-figures should be indicated. Each paper shall conclude with a critical summary not exceeding 350 words.

Drawings, diagrams or other illustrations should be made on larger scale (preferably) twice the size than the ones in which they are intended to appear. They should be done in Indian ink on bristol board with lettering in pencil. Scale of magnification of camera lucida tracings should be indicated by the side of drawings. In certain special cases arrangements will also be made for monochrome lithographic and other colour plates. Reduction of illustrations desired should be indicated in pencil. Appropriate legends should accompany all drawings. Names of authors are to be marked in pencil on the left-hand corner of drawing sheets. Photomicrographs should be securely mounted with colourless paste.

All tables, quotations and footnotes which will be set hereafter (beginning from Vol. I, No. 2) in types smaller than the text, should be typewritten on separate sheets and placed with the text in proper sequence. Footnotes should be numbered in Arabic numerals.

References to literature in the text should be given, whenever possible, in chronological order, only the names of authors and years of publication, in brackets, being given. They should be cited in full after the summary, the authors' names following in alphabetical order. Thus,

Name or Names of author; Name of Journal (abbreviation) with a single underline; Year of publication; Number of Volume with a double underline, and lastly page. The following would be a useful illustration:—

Bergmann and Stather . . . . . Z. Physiol. Chem., 1926, 152, 189.

Two copies of slip-proof and wherever possible, a page proof for final revision will be sent to authors. All corrections are best made on the slip-proof which should be transmitted to the Office of the Academy. All proof corrections involve heavy expenses which would be negligible if the papers are carefully revised by the authors before submission.

Authors will be supplied with reprints of their papers at cost price, if intimation is given at the time of returning the slip-proof.

Blocks appearing in the *Proceedings* will be available for purchase by their respective authors. Orders for the same should be sent along with the corrected proofs and in any case not later than one month after the date of publication of the paper. The price charged would be 25% of the actual cost of the blocks *plus* freight and despatching charges. If the blocks are reproduced in other journals or publications, due acknowledgment should be made in them to the *Proceedings*.

The original drawings and plates of blocks appearing in the *Proceedings* will be returned to such of the authors as may require them provided the cost of despatching such originals is borne by them.

CONTENTS

PAGE

The Structure and Optical Behaviour of Jadeite . . . . . Sir C. V. Raman and A. Jayaraman 117

Raman Spectrum of Crystalline Cadmium Sulphate . . . . . T. S. Krishnan and P. S. Narayanan 121

Relaxation Effects in Aqueous Solutions of Sulphur Dioxide . . . . . M. Krishnamurthi 125

Achromatic Combinations of Birefringent Plates. Part I. An Achromatic Circular Polarizer . . . . . S. Pancharatnam 130

Achromatic Combinations of Birefringent Plates. Part II. An Achromatic Quarter-Wave Plate . . . . . S. Pancharatnam 137

Inverse Probability and Evolutionary Markoff Stochastic Processes . . . . . Alladi Ramakrishnan 145

Scattering Measurements in Nuclear Emulsions and Their Application to Measuring the Charge of Primary Cosmic Ray Nuclei . . . . . S. Biswas, B. Peters and Rama 154